

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
19 February 2004 (19.02.2004)

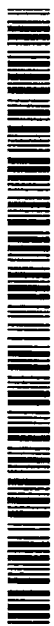
PCT

(10) International Publication Number
WO 2004/015139 A1

- (51) International Patent Classification⁷: C12Q 1/68 (74) Agent: SCHUBERT, Klemens; Neue Promenade 5, 10178 Berlin-Mitte (DE).
- (21) International Application Number: PCT/EP2003/008602 (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 1 August 2003 (01.08.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
102 36 406.0 2 August 2002 (02.08.2002) DE
- (71) Applicant (*for all designated States except US*): EPIGENOMICS AG [DE/DE]; Kastanienallee 24, 10435 Berlin (DE).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): ADORJAN, Peter [HU/DE]; Dunckerstrasse 4, 10437 Berlin (DE). PIEPENBROCK, Christian [DE/DE]; Schwartzkoffstrasse 7 B, 10115 Berlin (DE). RUJAN, Tamas [HU/DE]; Hiddenseer Strasse 13, 10437 Berlin (DE). SCHMITT, Armin [DE/DE]; Hortensienstrasse 29, 12203 Berlin (DE).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHOD FOR AMPLIFICATION OF NUCLEIC ACIDS OF LOW COMPLEXITY

(57) Abstract: The invention describes a method for amplifying nucleic acids, such as DNA with means of an enzymatic amplification step, such as a polymerase chain reaction, specified for template nucleic acids of low complexity, e.g. pre-treated DNA, like but not limited to DNA pre-treated with bisulfite is disclosed. The invention is based on the use of specific oligo-nucleotide primer molecules to solely amplify specific pieces of DNA. It is disclosed how to optimize the primer design for a PCR if the template DNA is of low complexity.



WO 2004/015139 A1

**Method for amplification of nucleic
acids of low complexity**

5 The present invention relates to a method for the amplification of nucleic acids.

Description

10 This invention relates to the fields of genetic engineering, molecular biology and computer science, and more specifically to the field of nucleic acid analysis based on specific nucleic acid amplification.

15 The matter of the present invention is a method for amplifying nucleic acids, such as DNA by means of an enzymatic amplification step, such as a polymerase chain reaction, specified for template nucleic acids of low complexity, e.g. pre-treated DNA, like but not limited to
20 DNA pre-treated with bisulfite. The invention is based on the use of specific oligo-nucleotide primer molecules to solely amplify specific pieces of DNA. It is disclosed how to optimize the primer design for a PCR if the template DNA is of unusually low complexity. Also, for the
25 optimal primer design it was considered that the treated template DNA is single stranded.

 The amplification of nucleic acids relies mainly on a method called polymerase chain reaction (PCR). The PCR is
30 based on the activity of the enzyme DNA polymerase, which is elongating primer molecules, which bind to the template DNA by adding dNTPs and hereby copying the template sequence (Saiki RK, Gelfand DH, Stoeffel S, Scharf SJ, Higuchi R, Horn T, Mullis KB and Erlich HA (1988).
35 Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239 : 487-491). The

primer molecules are designed to specifically hybridize to those regions of the template DNA that define both ends of the amplificate. The forward primer binds to the 5' end of the sense strand of the amplificate, whereas the reverse primer binds to the 5' end of the reverse strand, hereby defining the starting points of the polymerase reaction and eventually determining the length of the amplificate.

Before the polymerase starts the template DNA gets denatured, this is usually done by a short cycle of heating the reaction mixture up to about 95°C, then cooling it down to the annealing temperature determined by the melting temperature of the primer molecules used and finally allowing the polymerase to elongate the annealed primers at its ideal working temperature for some minutes. This cycle is repeated several times each starting with the denaturation step. The primer molecules hybridize to the single stranded DNA. The forward primer is the starting molecule for a copy of the sense strand and the reverse primer is the starting molecule for a copy of the anti-sense strand.

These first copies will be of unspecific length, limited only by the polymerase's activity. However in the following cycle, the forward primer will also bind to the first copy of the anti-sense strand, the polymerase will take that copy as a template and will elongate the primer only as far as there is template DNA. Hereby the length of the second copy gets limited to the length defined by the first nucleotide of the second primer. In the following cycles more and more pieces of template DNA compete for the primer molecules and eventually the DNA amplificate of defined length will be the main product.

However, in the case of a bisulfite treated DNA the template DNA is single stranded. The bisulfite or similar treatment alters the original sequences on both strands such that these are not complementary to each other after the treatment. As a result no complementary strand to the target sequence exists. A first primer molecule binds to the one end of the single stranded target sequence. The polymerase elongates said primer and copies said target sequence. The second primer molecule cannot bind to the complementary, so called anti-sense strand, as it would in a standard PCR. Therefore the second primer molecule is designed to bind to the first copied sequence instead. More specifically it will bind to that part of the copied nucleic acid which is the complement to the other end of said target sequence.

The results of a PCR are highly depending on the choice of the ideal primer. The choice of a primer molecule must respect constraints permitting a correct amplification by PCR, fulfilling hybridization temperature conditions and auto- or hetero-hybridization prevention.

In other words, as any PCR requires two primer molecules to amplify a specific piece of DNA in one reaction the melting temperatures of both primers need to be very similar in order to allow proper binding of both at the same hybridization temperature. That is why most primer design programs require the user to define a preferred melting temperature or a permitted range of melting temperatures. This requirement becomes the limiting factor when designing primers for a so called multiplex PCR, as all primer pairs in use need to have the same or at least very similar melting temperatures. Additionally primers have to be very specific, in order to only amplify those pieces of DNA that are the target.

By providing the means for designing extremely accurate primer pairs for DNA hybridization procedures this invention relates to the so called PCR primer design. More specifically the body of this invention relates to the specific requirements of primers and therefore of primer design when using template DNA that consists of essentially only three different nucleotides and is single stranded. This is the case when using bisulfite treated DNA as a template, as it contains no cytosine other than the methylated cytosines in a CG dinucleotide and a rest of insufficiently treated and therefore untransformed non-methylated cytosines. The invention relates specifically to the primer design when using bisulfite treated DNA as template.

It would be obvious to an individual skilled in the art that the use of the primers as specified in this invention are not limited to nucleic acid amplification. Said primers can be used for several purposes, such as amplification, but also for nucleic acid sequencing or as blocking oligonucleotides during analysis of bisulfite treated DNA. Therefore the use of said primers is not limited to nucleic acid amplification but extends to all standard molecular biological methods.

Pairs of these primers are used to specifically amplify DNA from a small amount of sample DNA that consists of bisulfite treated DNA originating from a limited source of DNA like a bodily fluid or tissue sample.

DNA can occur methylated or non-methylated at certain positions and this information is relevant for the status of a genes transcription. The methyl group is attached to the cytosine bases in CpG positions. The identification of 5-methylcytosine in a DNA sequence as opposed to unmethylated cytosine is of greatest importance for example

- when studying the role of DNA methylation in tumorigenesis. But, because the 5-Methylcytosine behaves just as a cytosine for what concerns its hybridization preference (a property relied upon for sequence analysis) its positions can not be identified by a normal sequencing reaction. Furthermore in a PCR amplification this relevant epigenetic information, methylated cytosine or unmethylated cytosine, will be lost completely.
- 10 This problem is usually solved by treating the genomic DNA with a chemical leading to a conversion of the cytosine bases, which consequently allows to differentiate the bases afterwards.
- 15 A tool most useful for analyzing DNA methylation is the bisulfite conversion of DNA that converts cytosine bases into bases showing a hybridization behavior as thymine bases. Hereby the DNAs complexity is reduced by a fourth.
- 20 Bisulfite conversion is the most frequently used method for analyzing DNA for 5-methylcytosine. It is based upon the specific reaction of bisulfite with cytosine which, upon subsequent alkaline hydrolysis, is converted to uracil, whereas 5-methylcytosine remains unmodified under
- 25 these conditions (Shapiro et al. (1970) Nature 227: 1047). However, in its base pairing behavior, uracil corresponds to thymine, that is, it hybridizes to adenine; whereas 5-methylcytosine doesn't change its chemical properties under this treatment and therefore still has
- 30 the base pairing behavior of a cytosine, that is hybridizing with guanine. Consequently, the original DNA is converted in such a manner that methyl-cytosine, which originally could not be distinguished from cytosine by its hybridization behavior, can now be detected as the
- 35 only remaining cytosine using "normal" molecular biological techniques, for example, by amplification and hy-

bridization or sequencing. All of these techniques are based on base pairing which can now be fully exploited. Comparing the sequences of the DNA prior to and after bisulfite treatment allows an easy identification of those bases that have been methylated.

In the scope of this invention when it says "a nucleotide (...) was converted by the treatment..." this conversion is meant to be able to differentiate between methylated and un-methylated cytosine bases within said sample, as for example the conversion of un-methylated cytosine bases to bases which hybridize to adenine by the treatment with bisulfite.

An alternative method is to use restriction enzymes that are capable of differentiating between methylated and un-methylated DNA, but this is restricted in its uses due to the selectivity of the restriction enzyme towards a specific sequence.

An overview of the further known methods of detecting 5-methylcytosine may be gathered from the following review article: Rein T, DePamphilis ML, Zorbas H, Nucleic Acids Res. 1998, 26, 2255.

In terms of sensitivity, the prior art is defined by a method, which encloses the DNA to be analyzed in an agarose matrix, thus preventing the diffusion and renaturation of the DNA (bisulfite reacts with single-stranded DNA only), and which replaces all precipitation and purification steps with fast dialysis (Olek A, Oswald J, Walter J (1996) A modified and improved method for bisulfite based cytosine methylation analysis. Nucleic Acids Res. 24: 5064-6). Using this method, it is possible to analyze individual cells, which illustrates the potential of the method.

To date, barring few exceptions (e.g., Zeschnigk M, Lich C, Buiting K, Doerfler W, Horsthemke B (1997) A single-tube PCR test for the diagnosis of Angelman and Prader-Willi syndrome based on allelic methylation differences at the SNRPN locus. *Eur J Hum Genet.* 5: 94-8) the bisulfite technique is only used in research. Always, however, short, specific fragments of a known gene are amplified subsequent to a bisulfite treatment and either completely sequenced (Olek A, Walter J (1997) The pre-implantation ontogeny of the H19 methylation imprint. *Nat Genet.* 3: 275-6) or individual cytosine positions are detected by a primer extension reaction (Gonzalvo ML and Jones PA (1997) Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.* 25: 2529-31; WO 95/00669) or by enzymatic digestion (Xiong Z, Laird PW (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* 25: 2532-4).

Another technique to detect hypermethylation is the so called methylation specific PCR (MSP) (Herman JG, Graff JR, Myohanen S, Nelkin BD and Baylin SB (1996), Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci USA.* 93: 9821-6). The technique is based on the use of primers that differentiate between a methylated and a non-methylated sequence if applied after bisulfite treatment of said DNA sequence. The primer either contains a guanine at the position corresponding to the cytosine in which case it will after bisulfite treatment only bind if the position was methylated. Or the primer contains an adenine at the corresponding cytosine position and therefore only binds to said DNA sequence after bisulfite treatment if the cytosine was unmethylated and has hence been altered by the bisulfite treatment so that it hybridizes to adenine.

With the use of these primers amplicons can be produced specifically depending on the methylation status of a certain cytosine and will as such indicate its methylation state. The present invention, however, does preferably not include CpGs in the primer sequence.

Another new technique is the detection of methylation via Taqman PCR, also known as MethyLight (WO 00/70090). With this technique it became feasible to determine the methylation state of single or of several positions directly during PCR, without having to analyze the PCR products in an additional step.

In addition, detection by hybridization has also been described (WO 99/28498).

Further publications dealing with the use of the bisulfite technique for methylation detection in individual genes are:

Grigg G, Clark S (1994) Sequencing 5-methylcytosine residues in genomic DNA. *Bioassays* 16: 431-6; Zeschnigk M, Schmitz B, Dittrich B, Buiting K, Horsthemke B, Doerfler W (1997) Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. *Hum Mol Genet.* 6: 387-95; Feil R, Charlton J, Bird AP, Walter J, Reik W (1994) Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing. *Nucleic Acids Res.* 22: 695-6; Martin V, Ribieras S, Song-Wang X, Rio MC, Dante R (1995) Genomic sequencing indicates a correlation between DNA hypomethylation in the 5' region of the pS2 gene and its expression in human breast cancer cell lines. *Gene* 157 : 261-4; WO 97/46705; WO 95/15373; WO 97/45560

35

For all those methods mentioned above, which are based on PCR amplification of bisulfite treated DNA, the biggest challenge is to design primers that are specific.

5

THE PROBLEM AND ITS SOLUTION

There are a number of programs available on the market that offer to design primer pairs in order to amplify a piece of DNA in a PCR. Usually they require as input the
10 template DNA sequence, the preferred melting temperature T_M , the desired length of the amplificate and optionally the preferred length of the primer molecules.

However if a primer is required to bind specifically to
15 bisulfite treated DNA, the design of the primer molecule is especially difficult and those tools known in the art are not competent to design primers that lead to specific products. The following problems occur when dealing with bisulfite treated DNA instead of standard DNA:

20

First, the sequence complexity of the bisulfite treated genome is reduced dramatically. Complexity in this context is meant to be a measure for the similarity of a given sequence to a random or stochastic sequence; the
25 more complex a sequence is the more it is similar to a random sequence. A reduced complexity of the genome means there are less degrees of variation. Where there are essentially only three different nucleotides rather than four, the probability of a sequence to occur twice in a given length of sequence is much higher. For example, a
30 primer molecule of 20 nucleotides in length is likely to be unique in the human genome, if it is not part of a repeat sequence: The human genome is known to consist of about 3×10^9 bases. There are $4^{20} \approx 10^{12}$ different ways
35 to form sequences of a length of 20 nucleotides, assuming equidistribution of the bases, which makes multiple oc-

- 5 currences of a given 20-mer (oligonucleotide of 20 nucleotides) extremely unlikely. However since there are only $320 \approx 3 \times 10^9$ different 20-mers possible over a 3-letter alphabet, this multiple occurrence cannot be excluded. In addition a bisulfite treated sequence, enriched in thymine in the sense strand and enriched in adenine in the reverse complementary strand, will contain more repeats and regions of general low complexity.
- 10 Another way to enhance or guarantee uniqueness of primer and/or oligo molecules is to estimate their expected frequency in the genome based upon a Markov model of order n for the human genome or to check their uniqueness explicitly by counting their exact occurrence. The estimation
- 15 based upon the Markov model relies upon the determination of the probabilities of all 4^n n -mers (oligo molecules of n nucleotides) in the human genome or in all amplicates which are used in the hybridization and the conditional probabilities of all four bases given these n -mers. The
- 20 primer pairs will be constructed from forward and reverse oligos which lie within an appropriate distance to each other and which have minimal individual expected occurrence elsewhere in the genome.
- 25 A second challenge in primer design for bisulfite treated DNA is that the melting temperature T_M of a bisulfite DNA primer of a certain length is typically lower than the melting temperature T_M of a standard primer containing cytosines. This is due to the fact that every cytosine in
- 30 a bisulfite treated DNA is - after amplification by PCR - replaced by thymine. Cytosine binds its corresponding base guanine via three hydrogen bonds, whereas thymine binds its corresponding base adenine via two hydrogen bonds only, leading to a generally weaker binding, a
- 35 lower T_M .

A third problem arises from the fact that bisulfite treated sequences are not only lacking cytosines but are also thymine-rich. Thymine also hybridizes unspecifically with guanine. This makes mismatching (unspecific binding of a primer to a sequence not identical) of a primer designed for bisulfite treated DNA much more likely than mismatching of a standard primer consisting of four different nucleotides.

It is the aim of this invention to overcome these problems, which are specific for primer based amplification of bisulfite treated DNA.

For a so called "multiplex PCR" it becomes especially difficult to design primer pairs. This expression is used to describe an experiment in which several different pieces of DNA are amplified simultaneously, in one reaction vessel and at the same time. Obviously this saves a lot of effort and time and is as such a basic requirement for high throughput assays based on PCR amplification. An overview on the state of the art concerning multiplex PCR is given by Henegariu et al. (Henegariu O, Heerema NA, Dlouhy SR, Vance GH and Vogt PH (1997) Multiplex PCR: Critical Parameters and Step-by-Step Protocol. BioTechniques 23: 504-511), who offer a step-by-step protocol on how to tackle multiplex PCR problems. However, the possibility of a special primer design is not mentioned in this article.

To ensure that the multiplex PCR works and the multiple products are amplified indeed usually a gel electrophoresis of the reaction mixture is performed. The products get separated due to their different sizes. Unfortunately, the ability of agarose gel electrophoresis to distinguish the products is slightly limited. However, it is possible to test for different product sizes with the

means of a fragment analyzer, which is much more accurate and able to distinguish product sizes of one base difference. Hence different product sizes are no longer a requirement to be considered in the primer design for a
5 multiplex PCR.

In patent WO 01/94634 a method for a multiplex PCR using at least two primer pairs is described that consists of basically a two step amplification procedure wherein one
10 step is referred to as pre-amplification. After pre-amplification (by means of PCR) with a number of primer pairs the sample gets divided into as many portions as there are primer pairs. At least one (and preferably only one) of the previously used primer pairs is added. This
15 method doesn't relate in any way to the selection or design of primer molecules described herein.

In an article by Shuber et al. (Shuber AP, Grondin VJ and Klinger KW (1995) A simplified procedure for developing
20 multiplex PCRs. Genome Res 5 (5) : 488-493) regarding multiplex PCR, the authors suggest to use primers, which contain a 3' region complementary to sequence specific recognition sites and a 5' region of a defined length of 20 nucleotides each. The authors claim that they could
25 establish identical reaction conditions, cycling times and annealing temperatures for any PCR primer pair following those requirements.

In several recent papers successful multiplex PCRs have
30 been established. For example, Becker et al. have reported the development of a multiplex PCR reaction for the detection of multiple staphylococcal enterotoxin genes, which uses individual primer sets for each toxin gene (Becker K, Roth R and Peters G (1998) Rapid and specific
35 detection of toxigenic Staphylococcus aureus : use of two multiplex PCR enzyme immunoassays for amplifica-

tion and hybridization of staphylococcal enterotoxin genes, exfoliative toxin genes, and toxic shock syndrome toxin 1 gene. J. Clin. Microbiol. 36: 2548-2553). This has been developed even further by Monday and Bohach, by
5 increasing the number of primer pairs applied in one reaction up to about 10 in order to have one assay to amplify all of the characterized enterotoxin genes. This still required a unique established primer pair for the detection of every individual gene (Monday SR and Bohach
10 GA (1999) Use of multiplex PCR to detect classical and newly described pyrogenic toxin genes in staphylococcal isolates. J. Clin. Microbiol. 37: 3411-3414).

In another paper by Sharma et al. a method for a one-
15 vessel-multiplex PCR is described wherein each of six chosen primer pair consists of one identical universal forward primer, based on a highly conserved region of those genes of interest and one reverse primer, specific for each individual gene. As such the assay leads to a
20 rapid amplification of a family of genes, which all have a conserved region in common. It is designed to detect presence or absence of certain genes in an unknown mixture. No further information is given about the primer design, apart from saying that they were designed by a-
25 alignment of published DNA sequences. This is certainly not the only requirement though, as one big limitation of the method is the need of getting PCR products of different sizes in order to identify those in the end (Sharma NK, Rees CED and Dodd CER (2000) Development of a single-
30 reaction multiplex PCR toxin typing assay for Staphylococcus aureus strains. Applied and Environmental Microbiology 66 (4) : 1347-1353).

In the patent application WO 01/36669 a method is described
35 which uses a similar approach for the controllable amplification of a higher number of sequences in selec-

ting one randomly chosen reverse primer that hybridizes unspecifically and a number of specific forward primers to amplify a group of sequences. As the reverse primer is labeled all products formed will be labeled as well. By hybridizing said amplicons towards immobilized detection oligos, which are able to differentiate the products, it will be easy to see which products have been amplified and herein the presence or absence of said sequences in the mixture can be determined.

The big disadvantage in all these methods is that every primer pair needs to be established individually first to ensure that a PCR product of the expected size was produced and that no additional or nonspecific products are generated. Once the specificity of the primer pairs had been determined, PCR conditions, buffers, and primer concentrations need to be optimized to establish conditions under which the primer molecules can be combined into one single PCR reaction without affecting the ability of the primer pairs to generate a gene specific amplicon.

A more recently published approach by Nicodeme and Steyaert describes the conditions required for multiplex PCR and suggests an algorithm to automatically select for primer pairs (Nicodeme P and Steyaert JM (1997) Selecting optimal oligonucleotide primers for multiplex PCR. Proc. Int. Conf. Intell Syst Mol Biol; 5 : 210-213). In this approach the conditions for pre-selecting primer pairs for a successful one locus amplification (singleplex PCR conditions) are rather broad. The three basic requirements are the pairing distance between a forward and a reverse primer, the condition of non-palindromicity of a primer, and the condition that the 3' end of a primer must not be reverse complementary to any of the other primers sequence. This selection is done with the help of a typical primer design program called PRIMER. However,

PRIMER is a two step program, and in this approach the new method to design primers for a multiplex PCR takes the output from step 1 as input, which is a list of possible forward and a list of possible reverse primers for every amplificate.

The only further selection criteria for the multiplex PCR primers are the absence of the reverse complementarity of their 3' end towards the other primer sequences in the experiment. A second critical factor considered here is the GC versus AT ratio. To some extent it is this ratio that determines the melting temperature of a primer pair. The authors suggest to limit the GC/AT ratio to be inside a given range which would enable the simultaneous hybridization of several primer pairs at one reaction temperature. The final requirement is the electrophoresis distance, determined by the tool that is used to differentiate the PCR products in, for example, a gel electrophoresis. This most common method requires the products to be of different sizes. The whole concept of this method also requires to have a pool of possible primer pairs for each amplicon.

The design of suitable primers for a multiplex PCR on bisulfite treated DNA is an even greater challenge. The low complexity of the DNA, being reduced to essentially three different bases rather than four different bases, requires an extra careful selection of primers to avoid mismatching and unwanted amplification.

In the scope of this invention the word "mismatching" corresponds to the situation when the alignment of two sequences which are essentially complementary reveals positions in one of the sequences where the nucleotide base does not align with its corresponding base but a different one. The corresponding or complementary base pairs are adenine and thymine, cytosine and guanine,

are adenine and thymine, cytosine and guanine, uracil and adenine. For example, a cytosine that aligns with a thymine in its otherwise complementary sequence creates a mismatch of one base or nucleotide.

5

Accordingly "base mismatches" refers to the situation of a base mismatching with another as explained above, respectively "one or more base mismatches" refers to one or more bases (in a given sequence) that cannot be aligned with their corresponding bases.

10

Also, when the alignment reveals single nucleotide gaps in one of the aligned sequences this is understood under the term "mismatch" in the scope of this invention.

15

A 'gap' is to be understood as follows: If an alignment reveals that, in order to get the highest number of corresponding base pairs aligned, some bases are lacking a corresponding base in its otherwise complementary sequence, this is called a gap. Such a gap can have a length of one or more nucleotides.

20

To solve the problems mentioned above we invented a method consisting of several steps that is applicable for the amplification of nucleic acids in singleplex as well as in multiplex PCR experiments.

25

SUMMARY OF THE INVENTION

The method is comprised of the following steps:
Firstly, the nucleic acid sample containing the region of interest, which is to be amplified, is isolated. Secondly, this nucleic acid sample is treated in a manner that differentiates between methylated and un-methylated cytosine bases within said sample. Thirdly, a reaction mixture is set up containing a) the treated template nu-

30
35

cleic acids, carrying the region of interest (also called: target nucleic acid) that is to be amplified, b) specified oligo-nucleotide primers, c) an enzyme capable of amplifying said nucleic acids in a defined manner, d) 5 the necessary nucleotides required for the nucleic acid synthesis and e) a suitable buffer.

Said specified oligo-nucleotide primers are characterized in that
10 their sequences each reach a predefined measure of complexity (as described in detail below)
every possible combination of two primer molecules in said reaction mixture has a melting temperature below a specified threshold temperature
15 none of the possible combinations of two primer molecules in said reaction mixture leads to the amplification of an additional unwanted product as determined by virtual testing for amplification.

20 In the last step of the method said amplified target nucleic acid is detected by means commonly used by one skilled in the art.

The invention is composed of a method for the amplification of nucleic acids comprising the following steps of
25 isolating a nucleic acid sample, treating said sample in a manner that differentiates between methylated and unmethylated cytosine bases within said sample, amplifying at least one target sequence, within said treated nucleic
30 acid, by means of enzymatic amplification and a set of primer molecules, wherein said primer molecules are characterized in that
a) each primer molecule sequence reaches a predefined measure of complexity, b) every combination of any two
35 primer molecules in the set has a melting temperature below a specified threshold temperature and c) every combi-

nation of two primer molecules, under conditions allowing for one or more base mismatches per primer, does not lead to the amplification of an unwanted product when virtually tested using the treated and the untreated sample nucleic acids as template and the last step of detecting said amplified target nucleic acid.

More detailed description of the method:

The method is comprised of the following steps:
In the first step of the method, the nucleic acid sample, which contains the region of interest that is to be amplified, must be isolated from tissue or cellular sources. Such sources may include at least one cell, but usually several cells, cell lines, histological slides, bodily fluids, or tissue embedded in paraffin.

In a preferred embodiment of this invention the nucleic acid sample is isolated from a bodily fluid, a cell culture, a tissue sample or a combination thereof.

For example a certain kind of organ sample from a patient or an animal can be used to extract genomic DNA by the usually applied methods. Preferably, in this invention DNA is extracted from a tissue sample or a biological fluid like blood, serum, urine or other fluids. 'Bodily fluid' herein refers to a mixture of macromolecules obtained from an organism. This includes, but is not limited to, blood, blood plasma, blood serum, urine, sputum, ejaculate, semen, tears, sweat, saliva, lymph fluid, bronchial lavage, pleural effusion, peritoneal fluid, meningeal fluid, amniotic fluid, glandular fluid, fine needle aspirates, nipple aspirate fluid, spinal fluid, conjunctival fluid, vaginal fluid, duodenal juice, pancreatic juice, bile and cerebrospinal fluid. This also includes experimentally separated fractions of all of the

preceding. 'Bodily fluid' also includes solutions or mixtures containing homogenized solid material, such as feces.

5 The nucleic acids may include DNA or RNA. Isolation may be by means that are standard to one skilled in the art, this includes for example extraction of DNA with the use of detergent lysates, sonification and vortexing with glass beads. An example is the extraction of DNA from a
10 piece of a plant, like a leave or fruit. Once the nucleic acids, like genomic double stranded DNA, have been extracted they are used in the analysis.

15 In a preferred embodiment of this invention the nucleic acid sample is comprised of plasmid DNA, BACs (bacterial artificial chromosomes), YACs (yeast artificial chromosomes) or genomic DNA.

20 In another especially preferred embodiment of this invention the nucleic acid sample is comprised of human genomic DNA. It is preferred that the nucleic acids are of human origin.

25 In the second step, this nucleic acid sample is treated in a manner that differentiates between methylated and un-methylated cytosine bases within said sample. Cytosine bases which are unmethylated at the 5'-position are converted to uracil, thymine, or another base which is dissimilar to cytosine in terms of hybridization behavior.
30 This will be understood as 'treatment' hereinafter. The method most commonly used so far is the so called bisulfite treatment.

35 This step is of essential meaning to the process as it translates the methylation pattern of said nucleic acids into a pattern that is something like an imprint of the

methylation status itself. It contains essentially the same information but the pre-treated nucleic acids are no longer sensitive to amplification via PCR. Amplification via PCR does not differentiate between methylated and un-methylated cytosines and therefore leads to the loss of this level of information. The original methylation status however can be deducted whenever the described pre-treatment had been performed prior to the amplification step. Hence any means suitable to differentiate between a methylated and an un-methylated cytosine base are applicable, as long as the modified bases are still capable of being amplified by enzymatic means after treatment.

It is a preferred embodiment of this invention that said sample is treated by means of a solution of a bisulfite, hydrogen sulfite or disulfite. A treatment of genomic DNA as described above is carried out with bisulfite (hydrogen sulfite, disulfite) and subsequent alkaline hydrolysis which results in a conversion of non-methylated cytosine nucleobases to uracil or to another base which is dissimilar to cytosine in terms of base pairing behavior.

In the third step of this method, a reaction mixture is set up containing a) the treated template nucleic acids, comprising the region of interest (also called target nucleic acid) that is to be amplified, b) specified oligonucleotide primers, c) an enzyme capable of amplifying said nucleic acids in a defined manner, for example a polymerase, d) the necessary nucleotides required for the nucleic acid synthesis and e) a suitable buffer. The template nucleic acid contains at least one target nucleic acid, which is amplified in the reaction. One primer molecule of the at least one primer pair in the reaction mixture is capable of binding to the 3' end of one specified target nucleic acid. The first primer binds to the

3' end of the target sequence, this primer is elongated and a complementary sequence to the target sequence is made. The polymerase stops to elongate unspecifically. The next cycle starts by thermally denaturing the now
5 double stranded template nucleic acid into single stranded template nucleic acids. This is followed by the next phase of annealing when both primer molecules specifically bind to the target nucleic acid and its complementary strand. The second primer is identical to the 5'
10 end of the target molecule. It doesn't bind to the target sequence itself but to said complementary nucleic acid to the target sequence, as soon as this is denatured from the template.

15 The process is finished by the actual amplification phase at a slightly lower reaction temperature, during which the enzyme, for example the polymerase elongates the primer as a complementary sequence to the target nucleic acid. The polymerase elongates this second primer by using
20 the first copy as template until the end of said copied nucleic acid is reached. That way an identical copy to the original single stranded target nucleic acid is created. Hence, the length of the amplificate is determined by choosing the two primers.

25 The elongation products, being complementary to each other and hereby building a double stranded version of the target nucleic acid, serve as additional targets for the primer molecules binding in the next cycle of amplification.
30

Essentially step 3 of the method is comprised of amplifying at least one target sequence, within said treated nucleic acid, by means of enzymatic amplification and a set
35 of primer molecules.

Said primer molecules used in said method are characterized in that they, in addition to fulfilling all the usual requirements towards a PCR primer as will be specified in more detail later, also fulfill the following requirements:

Firstly, the sequence of each primer molecule used in step 3 of this method reaches a predefined measure of complexity.

10

In a preferred embodiment of this method the primer molecules are reaching a certain value of linguistic complexity. A notion and a measure of linguistic complexity has been introduced by Trifonov in 1990 and has been used for analysis of nucleotide sequences before (Trifonov, EN (1990) Making sense of the human genome. In Structure & Methods. Vol 1 pp 69-77 (eds. Sarma, RH and Sarma MH, Adenine Press, Albany, US). The linguistic complexity technique allows a calculation to be made of the structural complexity of any linear sequence of characters irrespective of whether the text is cognized or presently undeciphered. The sequences are compared exclusively from the point of view of their structural complexity with no reference to the meaning of the texts. In 1997 Trifonov published how the linguistic complexity of nucleosomal sequences is defined (Bolshoy, A; Shapiro, K; Trifonov, E and Ioshikhes I. (1997) Enhancement of the nucleosomal pattern in sequences of lower complexity. NAR 25 (16): 3248-3254). Quote: "The linguistic complexity measure exploits the major distinguishing feature between natural nucleotide sequences and uniformly random ones: the repetitiveness of the natural sequences, i.e. the frequent repetition, not necessarily a tandem one, of some oligonucleotides ("words"), while others are avoided. (...) Complexity can be directly calculated as the extent to which the maximal possible vocabulary (all word sizes

considered) is utilized in a given strength of sequence (...)."

5 In another preferred embodiment of this method said measure of complexity is set by the so called Shannon entropy (Shannon, C E, (1948) A Mathematical Theory of Communication, University of Illinois Press, Urbana). This is the most common measure to assess the information content (in a technical, non-semantic meaning) of linear information
10 carriers. It attributes the maximal value (which can be chosen to be 1 without restrictions) to sequences where all symbols (characters) occur at equal probability and a value of 0 to sequences consisting of just one repeated symbol (character, letter). A derived and more general
15 measure is the higher order Shannon entropy which attributes maximal value to sequences where all its subsequences occur at equal probability and a value of 0 or close to 0 to sequences consisting of periodic repetitions of short subsequences. The practical determination
20 of the (higher order) Shannon entropy however is limited by the finite lengths of sequences which often does not permit a precise estimation of the probability distribution of their constitutive symbols.

25 Further possible measures are for example the Lempel-Ziv complexity (Lempel, LB and Ziv, J (1976) On the complexity of finite sequences. IEEE Trans. Inf. Theory IT-22, 75-81) , the grammar complexity (Ebeling, W; Jimenez-Montano, MA (1980) On Grammars, Complexity and Information Measures of Biologicial Macromolecules. Mathematical
30 Bioscience 52, 53-71), the algorithmic complexity (Chaitin, 1990) and the conditial entropy.

35 Secondly, said primer molecules are also characterized in that every possible combination of any two primer molecules, in the set, has a melting temperature below a

specified threshold temperature. That way the accumulation of dimers caused by the binding of two primer molecules to each other in said reaction mixture is excluded. The number of primer pairs used in that step can be any
5 between one and n , leading to one or n amplificates respectively (n being a natural number).

As mentioned in the text the word "dimer" refers to a secondary structure formed by the hybridization of two
10 primer molecules to each other.

As referred to in the text 'melting temperature' refers to the temperature at which 50% of the nucleic acid molecules are in duplex and 50% are denatured under standard
15 reaction solution conditions.

Some primer design tools disqualify a primer if, besides the target sequence, a second identical sequence can be found in the template. However, due to the higher probability of a bisulfite primer to mismatch with non-
20 identical bisulfite treated DNA, it is an embodiment of this invention that only those primers are allowed to be used in said amplification method, for which no sequence homology can be found, to the extent that even those sequences that are different and/or mismatching in several
25 nucleotides are excluded. However, this would exclude primer molecules unnecessarily. Therefore they are only excluded if two primer molecules match to the template in a distance allowing for the amplification of an unwanted
30 product. This test is performed by means as, for example, the Electronic PCR. Electronic PCR (e-PCR) is an in silico virtual PCR carried out in order to assess the suitability of primer molecules prior to in vitro PCR. In the scope of this invention this testing will be called
35 'virtual testing' and it will be referred to as "virtually tested" or "virtually testing".

Thirdly, the primers used in step 3 of this invention are characterized, in that every possible combination of two primer molecules, in said reaction mixture, does not lead to the amplification of an additional unwanted product, when virtually testing for amplification using the treated and the untreated nucleic acid sample as template, even under conditions allowing for at least one base but not more than 20% of the total number of bases per sequence mismatching per primer. In the scope of this invention it is to be understood that those primer molecules are considered to bind to the template for which a template sequence exists that is in at least 80% of its nucleotide sequence identical to the target sequence the primer originally has been designed for. For example, a primer molecule of 50 nucleotides length is considered to still hybridize to a template sequence that differs in less than 11 nucleotides (= is identical in at least 80% of its nucleotide sequence) from the according target sequence. If a match is considered to be possible it has to be tested whether this match would lead to the amplification of an unwanted product. This can be done with the use of a program similar to e-PCR (see below).

Especially preferred is an embodiment of said method wherein the ability of said primer molecules to amplify an unwanted product is tested by means of in silico PCR, taking as template nucleic acid the coding strand of the treated sample, the non-coding strand of the treated sample and both of the strands of the untreated sample. It is especially preferred to perform the virtual testing with a tool like electronic PCR on the pretreated, preferably bisulfite treated, template sequence consisting of the treated sense and the treated anti-sense strand, and on the unconverted template.

Furthermore it is preferred that this treatment is bisulfite treatment and hence the nucleic acid template is the bisulfite converted coding strand of the human genome, the bisulfite converted non-coding strand of the human genome and both of the strands of the untreated human genome. Preferred is an embodiment of said method wherein the ability of said primer molecules to amplify an unwanted product is tested by means of electronic PCR, hereby taking as template nucleic acid the bisulfite converted coding strand of the human genome, the bisulfite converted non-coding strand of the human genome and both of the strands of the untreated human genome.

It is preferred that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than 20% of the number of nucleotides of the primer.

It is also preferred that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than 10% of the number of nucleotides of the primer.

It is especially preferred that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than 5% of the number of nucleotides of the primer.

It is a preferred embodiment of this invention that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than seven.

It is especially preferred that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than five.

5

It is another preferred embodiment of this invention that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than three.

10

It is especially preferred in the scope of this invention that the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is one.

15

It is also included in the scope of this invention to consider such primer molecules as being sufficiently similar to facilitate their binding to the template sequence, for which a template sequence can be found that differs in the number of nucleotides but is otherwise identical to the target sequence. When the alignment of the primer and the template sequence leads to a gap of up to 20% of the nucleotides of one sequence, preferably of the primer sequence, this shall still be considered to be sufficient for binding and hence potentially leading to the amplification of an unwanted product. Therefore these primers also need to be tested with the means of virtual PCR (for example with a program like e-PCR). Only if this test reveals the virtual amplification of an unwanted product caused by the combination of two primers, the according primer pairs are excluded from the set of selected pairs.

20

25

30

35

It is preferred that the number of nucleotides creating one gap, in one of the sequences, when aligning the primer molecule sequence with the template sequence, allowed

for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than 20% of the number of nucleotides of the primer molecule.

5

It is also preferred that the number of nucleotides creating one gap, in one of the sequences, when aligning the primer molecule sequence with the template sequence, allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than 10% of the number of nucleotides of the primer molecule.

10

It is preferred that the number of nucleotides creating one gap, in one of the sequences, when aligning the primer molecule sequence with the template sequence, allowed for when virtually testing the amplification of unwanted products according to step 3 c) of the invention is less than 5% of the number of nucleotides of the primer molecule.

15

20

Both of these situations, mismatching due to an alternative nucleotide or no-matching due to a missing nucleotide, are meant to be covered in the expression describing those primer molecules that will eventually be selected : "said primer molecules are characterized in that every combination of two primer molecules, under conditions allowing for one or more base mismatches per primer, does not lead to the amplification of an unwanted product when virtually tested using the treated and the untreated sample nucleic acids as template".

25

30

It is also preferred that the primer molecules that exceed a pre-specified melting temperature when binding to the template have to be virtually tested for amplification of unwanted products using the treated and the untreated

35

ted sample nucleic acids as template according to step 3
c) of the method.

5 The basic problem of finding a primer specific enough to
give only one product on the little complex bisulfite
DNA, is finally solved by testing each potential primer
pair for hybridization across the whole bisulfite con-
verted human genome. This requires translating the whole
10 human genome sequence information virtually into its bi-
sulfite treated version before performing a similarity
search against the primer pairs, which is based on a
method like the so called e-PCR (Schuler G.D. (1997) Se-
quence Mapping by electronic PCR. Genome Research 7(5):
541-550). However, as the bisulfite conversion results in
15 two no longer complementary strands this virtual hybridi-
zation test needs to be done against both bisulfite con-
verted strands. In addition in most cases the template
DNA is contaminated with unconverted genomic DNA. To also
exclude unwanted amplification on the unconverted DNA as
20 template, the same hybridization test has to be performed
a third time using the whole human genome sequence as a
template.

25 Therefore it is a preferred embodiment of this invention
that the ability of said primer molecules to amplify an
unwanted product is tested by means such as electronic
PCR.

30 In the last step of the method said amplified target nu-
cleic acid gets detected by any means standard to one
skilled in the art.

In a preferred embodiment of this method the set of
primer molecules is comprised of at least two primer
35 molecules but not more than 64 primer molecules, given

the number is a multiple of 2; in other words, the set is comprised of 1-32 primer pairs.

5 In another preferred embodiment of this method the set of primer molecules is comprised of between 2 and 32 primer molecules, given the number is a multiple of 2; in other words the set is comprised of 1-16 primer pairs.

10 In a preferred embodiment of this method, said primer molecule comprises at least one nucleotide within the last three nucleotides from the 3' end of the molecule, wherein said nucleotide is complementary to a nucleotide of the target sequence that, as a result of the treatment performed in step 2) of the invention, changed its
15 hybridization behavior.

It is a preferred embodiment of this method, that said primer molecule comprises at least one nucleotide within the last three nucleotides from its 3' end that is complementary to a nucleotide of the target sequence that
20 was converted by the treatment performed in step 2 of the method to another base exhibiting an alternative base pairing behavior.

25 In an especially preferred embodiment said nucleotide is a cytosine prior to the treatment that converts unmethylated cytosines. In a preferred embodiment said treatment is bisulfite treatment. Said primer molecule comprises at least one nucleotide within the last three nucleotides
30 from the 3' end of the molecule, wherein said nucleotide is complementary to a cytosine, that was converted by bisulfite treatment to another base exhibiting the base pairing behavior of thymine.

35 This is to exclude binding of said primer molecules to the remaining untreated or un-sufficiently treated nu-

cleic acids, which might still serve as template nucleic acid in the PCR.

5 Furthermore it is a preferred embodiment of this invention that said primer molecules do not form loops or hairpins on their own or with each other.

10 In another preferred embodiment of the method said primer molecules do not form dimers with each other.

15 In the text the word 'hairpin' is taken to mean a secondary structure formed by a primer molecule when the 3' terminal region of said nucleic acid hybridizes to the 5' terminal region of said nucleic acid forming a double stranded stem structure and wherein only the central region of the primer is single stranded.

20 As described in the text the word 'loop' refers to a secondary structure formed by a primer molecule when two or more nucleotides of said molecule hybridize thereby forming a secondary structure comprising a double stranded structure one or more base pairs in length and further comprising a single stranded region between said double stranded region.

25 The binding of a primer molecules 3' end to any part of a second primer molecule in the set needs to be avoided. Otherwise the polymerase would extend the first primer using the second primer as template, which would lead to
30 a new unwanted product, an extended primer, or rather a primer-hybrid, which would serve as the preferred template for the next round of the polymerase chain reaction and thereby prevent a sufficient amplification of the
35 wanted product.

Therefore it is another preferred embodiment of this method that each of said primer molecules is characterized in that the last at least 5 bases at the 3' end of said primer molecule are not complementary to the sequence of any other primer molecule in the set.

It is also preferred that said primer molecules do not bind to nucleic acids which prior to treatment of step 2 contained a 5'-CG-3' site. This would lead to a binding of the primers to bisulfite treated nucleic acids, specifically depending on their cytosines methylation status. A CG corresponding primer would bind to the treated methylated version only, whereas a primer corresponding to TG would bind to the treated unmethylated version of these nucleic acids only. It is therefore preferred that said primer molecules do not contain nucleic acid sequences complementary or identical to nucleic acid sequences which prior to treatment of step 2 contained a 5'-CG-3' site.

In a preferred embodiment of this method said primer molecules are of a specified size range.

It is especially preferred that these primers are comprised of 16-50 nucleotides.

In a preferred embodiment of this method said primer molecules do not comprise sequences that are complementary to regions of the target nucleic acids that contained specified restriction enzyme recognition sites prior to the treatment that altered the unmethylated cytosines base pairing behavior. It is preferred that said primers are complementary to target sequences which prior to the treatment performed in step 2 of the invention did not contain specified restriction enzyme recognition sites.

By selecting for the right primer molecules also the amplification sequence is determined. That is why it has to be taken into account to only use those primer molecules that lead to amplification of nucleic acids containing a reasonable high number of CpG sites to be analyzed. Due to the treatment of step 2 of this invention these CpG sites, depending on the methylation status of the cytosine, are converted and will therefore either appear as CG dinucleotides or as TG dinucleotides in the amplificate.

It is preferred that said primer molecules amplify regions of nucleic acids that prior to bisulfite treatment comprise of more than eight 5'-CG-3' sites also referred to as CG dinucleotides.

It is also preferred that said primer molecules amplify regions of nucleic acids that prior to bisulfite treatment comprise of more than six 5'-CG-3' sites also referred to as CG dinucleotides.

It is also preferred that said primer molecules amplify regions of nucleic acids that prior to bisulfite treatment comprise of more than four 5'-CG-3' sites also referred to as CG dinucleotides and finally it is especially preferred that said primer molecules amplify regions of nucleic acids that prior to bisulfite treatment comprise of more than two 5'-CG-3' sites also referred to as CG dinucleotides.

Said primer molecules lead to amplification within a specified size range.

It is a preferred embodiment of this sequence that said primer molecules lead to amplicates which are comprised of at least 50 bp but not more than 2000 bp.

5 Especially preferred are primer molecules that lead to amplicates which are comprised of at least 80 bp but not more than 1000 bp.

10 Furthermore a method is preferred wherein said primer molecules lead to amplicates of treated nucleic acids which prior to the treatment which altered the unmethylated cytosines base pairing behavior did not contain restriction enzyme recognition sites. Said primer molecules lead to amplicates that are amplified regions of the
15 treated nucleic acids which prior to the treatment performed in step 2) of the method did not contain specified restriction enzyme recognition sites.

A further subject of this invention is a method on how to
20 produce said primer molecules. The main step of producing a primer molecule is determining its sequence. In the following the phrase "primer design" will be used instead of primer production, whenever it is referred to the step of determining said specific primer sequences. Designing
25 primer molecules is a process which as such is well known to scientists skilled in the art. The programs usually used for this purpose are such as PRIMER3 or OSP (Rozen S and Skaletsky H (2000) PRIMER3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365-386; Hillier L and Green P (1991) OSP: A computer
30 program for choosing PCR and DNA sequencing primers. PCR Methods and Applications 1: 124-128). Other primer design systems (like described in EP-A 1136932) are often based on those commonly known programs.

35

An embodiment of this invention takes advantage of using a program like PRIMER3 first, to then add a number of steps that finally result in an advanced method of designing primers that are specifically useful for amplifying sequences of low complexity.

In the first step of this method for designing specific primer molecules for nucleic acids of low complexity, primer pairs that amplify single products are selected by applying standard tools of primer design known in the art, like for example the program PRIMER3 (Rozen, S and Skaletsky, H (2000) Methods Mol Biol 132: 365-386).

In the second step of the method said primer pairs are tested whether or not one of its primer molecules when hybridizing to any other primer molecule in the set exceeds a specified threshold melting temperature T_M . If this is the case the primer pair that comprises of said primer is excluded from the set of potentially combined pairs.

In the third step of the method the number of previously selected primer pairs, is reduced to a smaller number by implementing as new criteria a measure for the primer sequence's complexity. Primer pairs that consist of a primer molecule which does not meet said criteria are excluded.

The basic problem of finding a primer specific enough to give only one product on the little complex bisulfite DNA, is finally solved by testing each potential primer pair for hybridization across the whole bisulfite converted human genome. This requires translating the whole human genome sequence information virtually (as in "in silico") into its treated, for example bisulfite treated, version before performing a similarity search against the

primer pairs, which is based on a method like the so called e-PCR (Schuler G.D. (1997) Sequence Mapping by electronic PCR. Genome Research 7(5): 541-550). However, as the bisulfite conversion results in two different versions of the double helix whose sense and anti-sense strands are no longer mutually complementary, this in silico amplification needs to be performed on both bisulfite converted versions of the genome. In addition in most cases the template DNA is contaminated with unconverted genomic DNA. It cannot be excluded that single cytosines or longer runs of DNA remain unconverted or are only converted incompletely by the bisulfite treatment. To also exclude unwanted amplification of the unconverted DNA as template, the same hybridization test has to be performed a third time using the whole human genome sequence as a template.

As this is quite some effort and requires time (CPU time) this is the fourth and last step of this design method, that is absolved prior to the final testing in a "wet", lab based, experiment.

In addition to improve the specificity of said primer molecules the stringency of the selection criteria is increased: Some standard primer design tools disqualify a primer if in the template sequence, a second identical sequence, besides the target sequence, can be found. That way mispriming at rather stringent hybridization conditions is avoided. This mispriming would not necessarily lead to an additional unwanted product, but would lead to the dilution of the primer molecules available for amplification. This selection has been performed in step one already (for example by PRIMER3). However, due to the higher probability of a bisulfite primer molecule to mismatch with non-identical bisulfite treated DNA, there is still a chance for said primer molecules to misprime even

when up to 20% of the nucleotides of the primer sequence differ. Therefore it is claimed in this invention to only use primer molecules for which not even a weak sequence homology can be found. However, this would exclude primer molecules unnecessarily. Therefore they are only excluded if two primer molecules match to the template and amplify an unwanted product. This test is performed by means as, for example, the Electronic PCR. Electronic PCR (e-PCR) is an in silico virtual PCR carried out in order to assess the suitability of primers prior to in vitro PCR.

In the fourth step of the method on how to design these primers it is therefore tested whether there are any regions of the template nucleic acid, said template being comprised of the sense and the anti-sense strand of the treated and the untreated nucleic acids, that are identical in sequence with the primer molecule to more than 80% and if those primer molecules are able to amplify an unwanted product. If this is the case, the primer pair comprising said primer molecule is excluded from the selection.

The template nucleic acid is comprised of the treated template nucleic acid and the untreated template nucleic acid. The treated nucleic acid in itself is comprised of a two strands which after treatment are not complementary to each other anymore. This virtual testing for example can be performed as described by Gregory Schuler in his article (cited above) about sequence mapping by "Electronic PCR". The primer pairs remaining can be used to specifically amplify regions of nucleic acids of low complexity, which is the aim of this invention. Hence step 4 of the design method is the virtual testing of each possible primer pair combination, under pre-specified conditions at a stringency allowing for one or more base pair mismatches, as to whether no unwanted nucleic acids are

amplified. Said virtual testing is carried out upon both untreated and treated nucleic acids. The wording "possible combinations" refers to all combinations that are possible within a set of primer pairs to be used in one
5 amplification reaction vessel.

In a preferred embodiment an additional step is added following the virtual testing, which is testing in a lab based single PCR assay all those pairs that remained,
10 whether the desired amplificate can be obtained or not. If that is the case, the chosen pairs can be used to specifically amplify those regions of nucleic acids of low complexity according to the method as described before.

15 In a specially preferred embodiment the first step of the design method is characterized as selecting a pool of possible primer pairs per amplificate by means of a standard PCR primer design program using said nucleic acids as template that have been masked for repeats and SNPs
20 considering the following factors:
length of amplificate, length of primer, melting temperature of the primer molecule, dimer formation parameters, loop formation parameters, exclusion of unidentified or ambiguous nucleotides in the primer sequence, exclusion
25 of restriction enzyme recognition sites.

In a preferred embodiment of this invention this measure of complexity is a measure of linguistic complexity as defined by Bolshoy et al. (see above). Those primer pairs
30 are excluded from the previously selected ones, which comprise of one primer that doesn't reach a set level of this linguistic complexity.

In another preferred embodiment of this invention this
35 measure of complexity is a measure of Shannon entropy (as described before).

In an especially preferred embodiment of this design method, prior to performing step d) the additional step of excluding primer pairs from the remaining primer pairs
5 which consist of a primer molecule that comprises of at least one CpG site, is carried out.

In an especially preferred embodiment of this method according to the design of said primers, prior to performing step d) the additional step of excluding primer pairs from the remaining pairs when one of its primer molecules does not contain at least one nucleotide within the last three nucleotides from the 3' end of the molecule wherein said nucleotide is complementary to a nucleotide of the
10 target sequence that was converted to a different nucleotide by bisulfite treatment, is carried out.
15

In an especially preferred embodiment of this method according to the production of said primers, prior to performing step d) the additional step, of excluding primer pairs from the remaining primer pairs which amplify a nucleic acid that did not prior to treatment with bisulfite contain a minimum of two CpG sites, is carried out.
20

In an especially preferred embodiment of this method according to the production of said primers, prior to performing step d) the additional step of excluding primer pairs from the remaining primer pairs when one of its primer molecules contains more than 5 bases at its 3' end that are complementary to any other primer molecules sequence in the set, is carried out.
25
30

In an especially preferred embodiment of this method according to the production of said primers, prior to performing step d) the additional step of excluding from the remaining primer pairs those pairs, which comprise of one
35

primer molecule that in combination with another primer molecule in the set amplifies an unwanted product, when virtually testing according to step 3 c) of the amplification method under conditions allowing for a number of mismatching nucleotides of 20% of the number of nucleotides of the primer molecule, is carried out.

In an especially preferred embodiment of this method according to the production of said primers, prior to performing step d) the additional step of excluding from the remaining primer pairs those pairs, which comprise of one primer molecule that in combination with another primer molecule in the set amplifies an unwanted product, when virtually testing according to step 3 c) of the amplification method under conditions allowing for a number of nucleotides creating one gap, when aligning the primer molecule sequence with the template sequence, of up to 20% of the number of nucleotides of the primer molecule, is carried out.

In an especially preferred embodiment of this method according to the production of said primers, prior to performing step d) the additional step of excluding from the remaining primer pairs those pairs, which comprise of one primer molecule that in combination with another primer molecule in the set amplifies an unwanted product, when virtually testing according to step 3 c) of the amplification method under conditions allowing for four or less mismatching base pairs, is carried out.

In an especially preferred embodiment of this method according to the production of said primers, prior to performing step d) the additional step of excluding from the remaining primer pairs those pairs, which comprise of one primer molecule that in combination with another primer molecule in the set amplifies an unwanted product, when

virtually testing according to step 3 c) of the amplification method under conditions allowing for two or less mismatching base pairs, is carried out.

- 5 The following example is intended to illustrate the invention :

Example

- 10 Here we present experimental data that shows that multiplex PCRs designed with a tool according to this invention are more successful compared to multiplex PCRs not designed in this manner.
- 15 It is the aim of the experiment to amplify 40 different nucleic acids. The genomic regions of interest are given in the sequence protocol (SEQ ID 41-80). These genomic sequences were translated into their bisulfite converted versions and served as templates for amplification of
- 20 specific regions with the primer sequences described as follows.

- Primer molecule pairs used for single PCRs were originally designed with the use of the standard primer design
- 25 program PRIMER3 (as mentioned in the description). The criteria used in that step will not be discussed in detail. This selection however provides several possible primer pairs per amplificate. Following the present invention these primer pairs were selected further, according to the following criteria:
- 30

- The restriction enzyme recognition site to be excluded from the genomic nucleic acid (which subsequent to bisulfite conversion becomes the template
- 35 for the PCR amplification step) is : GTTTAAAC.

- The minimum length of the primer molecule is 18 nucleotides. The maximum length is 27 nucleotides. Ideally the primer consists of 22 nucleotides.
- 5 • The minimum required measure of linguistic complexity is 0.2.
- The minimum melting temperature of a primer molecule is 54°C and the maximum melting temperature is 57°C. The ideal melting temperature however is 55°C.
- 10 • The minimum length of an amplificate is 100 bp and the maximum length is 500 bp.
- 15 • The minimum number of CpG sites, that were present in the region of the nucleic acid, prior to bisulfite treatment, that was amplified is 4.
- The number of mismatch bases allowed for when virtually testing the primer pairs according to the invention for amplification of an unwanted product with the help of e-PCR (Electronic PCR) is 2.
- 20

25 The use of this invention, that is the use of either the design method, being the subject of the invention, and/or performing the steps of said method as described above (assuming a set size of 1) leads to the selection of the following 40 optimized primer molecule pairs:

30 TABLE 1:

| primer sequence | amplificate identifier | SEQ ID | number indicating primer direction | starting position of primer in the bisulfite converted sequence of the ROI |
|------------------------|------------------------|--------|------------------------------------|--|
| AATCCTCCAAATTCTAAAAACA | 2025 | 81 | 0 | 1816 |
| AGGAAAGGGAGTGAGAAAAT | 2025 | 82 | 1 | 2138 |

| primer sequence | amplificate identifier | SEQ ID | number indicating primer direction | starting position of primer in the bisulfite converted sequence of the ROI |
|---------------------------|------------------------|--------|------------------------------------|--|
| GGATAGGAGTTGGGATTAAGAT | 2044 | 83 | 0 | 2070 |
| AAATCTTTTCAACACCAAAT | 2044 | 84 | 1 | 2483 |
| AACCCTTTCTTCAAATTACAAA | 2045 | 85 | 0 | 1340 |
| TGATTGGGTTTTAGGGAAATA | 2045 | 86 | 1 | 1687 |
| TTGAAAATAAGAAAGGTTGAGG | 2106 | 87 | 0 | 1481 |
| CTTCTACCCCAAATCCCTA | 2106 | 88 | 1 | 1764 |
| TGTTTGGGATTGGGTAGG | 2166 | 89 | 0 | 2226 |
| CATAACCTTTACCTATCTCCTCA | 2166 | 90 | 1 | 2437 |
| TTTGTAGATTGAGGTTTTAGGGT | 2188 | 91 | 0 | 101 |
| ATCCATTCTACCTCCTTTTTCT | 2188 | 92 | 1 | 598 |
| GGAGGGGAGAGGGTTATG | 2191 | 93 | 0 | 133 |
| TACTATACACCCCCAAAACAA | 2191 | 94 | 1 | 506 |
| TTTTGGGAATGGGTTGTAT | 2194 | 95 | 0 | 1628 |
| CTACCCTTAACCTCCATCCTA | 2194 | 96 | 1 | 1996 |
| TTGTTGGGAGTTTTTAAGTTTT | 2212 | 97 | 0 | 1711 |
| CAAATTCTCCTTCCAAATAAAT | 2212 | 98 | 1 | 2063 |
| GTAATTGAAGAAAGTTGAGGG | 2267 | 99 | 0 | 1709 |
| CCAACAATAAACAACCTCT | 2267 | 100 | 1 | 2004 |
| GGAGTTGTATTGTTGGGAGA | 2317 | 101 | 0 | 1110 |
| TAAACCCCAATTTTCACTAA | 2317 | 102 | 1 | 1388 |
| TTTGTATTAGGTTGGAAGTGGT | 2383 | 103 | 0 | 1 |
| CCCAAATAAATCAACAACAACA | 2383 | 104 | 1 | 285 |
| GATTTTTGGAGAGGAAGTTAAG | 2387 | 105 | 0 | 789 |
| AAAATAAAACCAAACCCATA | 2387 | 106 | 1 | 1169 |
| TGGGGTTAGTTTAGGATAGG | 2391 | 107 | 0 | 1353 |
| CTTAAAAACACTAAACTTCTCAAA | 2391 | 108 | 1 | 1750 |
| TTTTTGTATTGGGGTAGGTTT | 2395 | 109 | 0 | 547 |
| CCCAACTATCTCTCCTCTATAA | 2395 | 110 | 1 | 1094 |
| ATTAGAAGTGAAAGTAATGGAATTT | 2401 | 111 | 0 | 381 |
| TCAATTTCCAAAAACCAAC | 2401 | 112 | 1 | 795 |

| primer sequence | amplificate identifier | SEQ ID | number indicating primer direction | starting position of primer in the bisulfite converted sequence of the ROI |
|--------------------------|------------------------|--------|------------------------------------|--|
| GGGATGGGTTATTAGTTGTAAA | 2453 | 113 | 0 | 1867 |
| CCTTCACACAAAACCTACAAAAA | 2453 | 114 | 1 | 2139 |
| TAATTGAAGGGGTTAATAGTGG | 2484 | 115 | 0 | 1861 |
| AAAACCAAAACCAAACTAAAA | 2484 | 116 | 1 | 2252 |
| AGTGGATTTGGAGTTTAGATGT | 2512 | 117 | 0 | 1016 |
| AACAAAATAAAAACTTCTCCCA | 2512 | 118 | 1 | 1446 |
| TAGGGGAAAAGTTAGAGTTGAG | 2741 | 119 | 0 | 1413 |
| CCCATTAACCCACAAAAA | 2741 | 120 | 1 | 1888 |
| ATTTTAGTTTGTGAAATGGGAT | 2745 | 121 | 0 | 1685 |
| TCTTAACCAATAACCCCTCAC | 2745 | 122 | 1 | 2097 |
| GTGGGTTTTGGGTAGTTATAGA | 2746 | 123 | 0 | 1679 |
| TAACCTCCTCTCCTTACCAA | 2746 | 124 | 1 | 2163 |
| TAGGATGGGGAGAGTAATGTTT | 2747 | 125 | 0 | 972 |
| ACAACTTATCCAACCTCCATTC | 2747 | 126 | 1 | 1448 |
| TCCCACAAAACTAAACAATTA | 2749 | 127 | 0 | 1370 |
| AGGTTTTAGATGAAGGGGTTT | 2749 | 128 | 1 | 1789 |
| TTTGAGGGGTTTAGTAGAAGTTA | 2751 | 129 | 0 | 88 |
| CCCAATAATCACAAAATAAACA | 2751 | 130 | 1 | 567 |
| ATACAACCTCAAATCCTATCCA | 2752 | 131 | 0 | 228 |
| AGGGAGAAGGAAGTTATTTGTT | 2752 | 132 | 1 | 712 |
| GGAAGATGAGGAAGTTGATTAG | 2755 | 133 | 0 | 1000 |
| CCTACAACCCTATCCTCTAAAA | 2755 | 134 | 1 | 1371 |
| TTAGTAGGGGTGTGAGTGTTTT | 2831 | 135 | 0 | 1313 |
| CAAACAAAACCTTCTATCTCAACC | 2831 | 136 | 1 | 1499 |
| TTATAGGGTTGAGTTTGGGAT | 2850 | 137 | 0 | 2100 |
| TAAACAAACAACAAATCTTCCA | 2850 | 138 | 1 | 2400 |
| TGAAAATGAAGGTATGGAGTTT | 2852 | 139 | 0 | 1262 |
| TTAAAACCATATAATCCCTCCA | 2852 | 140 | 1 | 1583 |
| TATGTTTGGTTTTGTTTTGAGA | 2859 | 141 | 0 | 1093 |
| AACCCCATCACTTTTATTTCTT | 2859 | 142 | 1 | 1491 |

| primer sequence | amplificate identifier | SEQ ID | number indicating primer direction | starting position of primer in the bisulfite converted sequence of the ROI |
|--------------------------|------------------------|--------|------------------------------------|--|
| GGGTGTAGAAGTGTTTAGGTTT | 2861 | 143 | 0 | 2385 |
| TTTCTCCCCTTACAACAATAAC | 2861 | 144 | 1 | 2732 |
| TCCCCTTCCAATATATCTCTC | 2864 | 145 | 0 | 884 |
| TGAGAGTGTTTTAGGGAAGTTT | 2864 | 146 | 1 | 1175 |
| AAAACCAAAACATAAACCAAAA | 2867 | 147 | 0 | 1312 |
| GATTAGGAGGGTTTGTTGAGAT | 2867 | 148 | 1 | 1701 |
| AATGGTTGATGATTTTGGTTT | 2961 | 149 | 0 | 2039 |
| ACTCTCTTCCCTATACCCCTAA | 2961 | 150 | 1 | 2311 |
| AGTTAGAAGAGGAGTTAGGATGG | 3511 | 151 | 0 | 1340 |
| TAATTTTCCAATACCCATTTTC | 3511 | 152 | 1 | 1711 |
| TGTTAGTAGAGTTTTAGGGAGGTT | 3532 | 153 | 0 | 1135 |
| ACACTACCTATCCTTACCCAC | 3532 | 154 | 1 | 1592 |
| TTTTTGTTTTTATGGGGTGTAT | 3534 | 155 | 0 | 1909 |
| TTAAATATCCCTTCCTTAACCA | 3534 | 156 | 1 | 2385 |
| TGGGTAGTATTTTTGTTGGTTT | 3538 | 157 | 0 | 956 |
| CCTAAAACTCTCTCATCCTCA | 3538 | 158 | 1 | 1414 |
| AGTGGTTTAGGAGTATTTGGTTA | 3540 | 159 | 0 | 659 |
| AACTCCCTCCATCTACAATATC | 3540 | 160 | 1 | 1064 |

These primer pairs lead to the amplification of specific regions (amplificates Seq IDs 1- 40) of the bisulfite converted sequences of the genomic ROIs (Seq IDs 41- 80) of interest. The ROIs can be identified by the four digit number that specifies the ROI and the corresponding amplificate - as indicated in the following table.

10 TABLE 2:

| SEQ ID | Class | Identifier | Kind of DNA | SEQ ID | Class | Identifier | Kind of DNA |
|--------|-------------|------------|--------------------|--------|-------|------------|------------------|
| 1 | amplificate | 2025 | bisulfite sequence | 41 | ROI | 2025 | genomic sequence |

| SEQ ID | Class | Identifier | Kind of DNA | SEQ ID | Class | Identifier | Kind of DNA |
|--------|-------------|------------|-------------------------|--------|-------|------------|-----------------------|
| 2 | amplificate | 2044 | bisulfite se- quence | 42 | ROI | 2044 | genomic se- quence |
| 3 | amplificate | 2045 | bisulfite se- quence | 43 | ROI | 2045 | genomic se- quence |
| 4 | amplificate | 2106 | bisulfite se- quence | 44 | ROI | 2106 | genomic se- quence |
| 5 | amplificate | 2166 | bisulfite se- quence | 45 | ROI | 2166 | genomic se- quence |
| 6 | amplificate | 2188 | bisulfite se- quence | 46 | ROI | 2188 | genomic se- quence |
| 7 | amplificate | 2191 | bisulfite se- quence | 47 | ROI | 2191 | genomic se- quence |
| 8 | amplificate | 2194 | bisulfite se- quence | 48 | ROI | 2194 | genomic se- quence |
| 9 | amplificate | 2212 | bisulfite se- quence | 49 | ROI | 2212 | genomic se- quence |
| 10 | amplificate | 2267 | bisulfite se- quence | 50 | ROI | 2267 | genomic se- quence |
| 11 | amplificate | 2317 | bisulfite se- quence | 51 | ROI | 2317 | genomic se- quence |
| 12 | amplificate | 2383 | bisulfite se- quence | 52 | ROI | 2383 | genomic se- quence |
| 13 | amplificate | 2387 | bisulfite se- quence | 53 | ROI | 2387 | genomic se- quence |
| 14 | amplificate | 2391 | bisulfite se- quence | 54 | ROI | 2391 | genomic se- quence |
| 15 | amplificate | 2395 | bisulfite se- quence | 55 | ROI | 2395 | genomic se- quence |
| 16 | amplificate | 2401 | bisulfite se- quence | 56 | ROI | 2401 | genomic se- quence |
| 17 | amplificate | 2453 | bisulfite se- quence | 57 | ROI | 2453 | genomic se- quence |

| SEQ ID | Class | Identifier | Kind of DNA | SEQ ID | Class | Identifier | Kind of DNA |
|--------|-------------|------------|-------------------------|--------|-------|------------|-----------------------|
| 18 | amplificate | 2484 | bisulfite se- quence | 58 | ROI | 2484 | genomic se- quence |
| 19 | amplificate | 2512 | bisulfite se- quence | 59 | ROI | 2512 | genomic se- quence |
| 20 | amplificate | 2741 | bisulfite se- quence | 60 | ROI | 2741 | genomic se- quence |
| 21 | amplificate | 2745 | bisulfite se- quence | 61 | ROI | 2745 | genomic se- quence |
| 22 | amplificate | 2746 | bisulfite se- quence | 62 | ROI | 2746 | genomic se- quence |
| 23 | amplificate | 2747 | bisulfite se- quence | 63 | ROI | 2747 | genomic se- quence |
| 24 | amplificate | 2749 | bisulfite se- quence | 64 | ROI | 2749 | genomic se- quence |
| 25 | amplificate | 2751 | bisulfite se- quence | 65 | ROI | 2751 | genomic se- quence |
| 26 | amplificate | 2752 | bisulfite se- quence | 66 | ROI | 2752 | genomic se- quence |
| 27 | amplificate | 2755 | bisulfite se- quence | 67 | ROI | 2755 | genomic se- quence |
| 28 | amplificate | 2831 | bisulfite se- quence | 68 | ROI | 2831 | genomic se- quence |
| 29 | amplificate | 2850 | bisulfite se- quence | 69 | ROI | 2850 | genomic se- quence |
| 30 | amplificate | 2852 | bisulfite se- quence | 70 | ROI | 2852 | genomic se- quence |
| 31 | amplificate | 2859 | bisulfite se- quence | 71 | ROI | 2859 | genomic se- quence |
| 32 | amplificate | 2861 | bisulfite se- quence | 72 | ROI | 2861 | genomic se- quence |
| 33 | amplificate | 2864 | bisulfite se- quence | 73 | ROI | 2864 | genomic se- quence |

| SEQ ID | Class | Identifier | Kind of DNA | SEQ ID | Class | Identifier | Kind of DNA |
|--------|-------------|------------|-------------------------|--------|-------|------------|-----------------------|
| 34 | amplificate | 2867 | bisulfite se- quence | 74 | ROI | 2867 | genomic se- quence |
| 35 | amplificate | 2961 | bisulfite se- quence | 75 | ROI | 2961 | genomic se- quence |
| 36 | amplificate | 3511 | bisulfite se- quence | 76 | ROI | 3511 | genomic se- quence |
| 37 | amplificate | 3532 | bisulfite se- quence | 77 | ROI | 3532 | genomic se- quence |
| 38 | amplificate | 3534 | bisulfite se- quence | 78 | ROI | 3534 | genomic se- quence |
| 39 | amplificate | 3538 | bisulfite se- quence | 79 | ROI | 3538 | genomic se- quence |
| 40 | amplificate | 3540 | bisulfite se- quence | 80 | ROI | 3540 | genomic se- quence |

5 The second task in this example is to select from these
40 primer pairs those pairs which can be combined in five
multiplex PCRs to amplify eight targets simultaneously.

10 The following steps, as disclosed in the invention, are
performed for selection of those subsets:

- 15 • The melting temperature of any combination of two of
those primer molecules hybridizing to each other
taking part in one multiplex experiment must be be-
low 20°C.
- The last seven nucleotides from the 3' end of every
primer molecule in a subset is used to check if
those are complementary and/or binding to any other
primer molecules' sequence used in the set.

20

- The number of mismatch bases allowed for when virtually testing the primer pairs for amplification of an unwanted product is 2. For this step every possible combination of 16 primer molecules in one subset is checked for its ability to amplify an unwanted product. This is done by means of e-PCR (electronic PCR).
- 10 Having performed all these steps results in the selection of three different optimized sets of primer molecule pairs that can be used in multiplex PCRs. These sets are in the following described as a set of numbers. Each number refers to a specific amplificate and therefore also
- 15 to a single primer pair (out of the list given above) which proved to be able to specifically amplify said nucleic acid in a single PCR experiment.

TABLE 3:

| optimized set 1 | | | | | | | | |
|-----------------|------|------|------|------|------|------|------|------|
| 8plex1 | 2194 | 2191 | 2391 | 2025 | 2961 | 3540 | 2861 | 2188 |
| 8plex2 | 2484 | 2106 | 2401 | 2850 | 3532 | 2044 | 2512 | 2852 |
| 8plex3 | 2453 | 2741 | 2867 | 2755 | 2267 | 2387 | 2864 | 2317 |
| 8plex4 | 2859 | 2383 | 2752 | 2747 | 2751 | 3511 | 2212 | 2746 |
| 8plex5 | 3534 | 2395 | 2745 | 3538 | 2749 | 2166 | 2831 | 2045 |
| optimized set 2 | | | | | | | | |
| 8plex1 | 2166 | 2212 | 3511 | 2383 | 2745 | 2859 | 3534 | 2861 |
| 8plex2 | 2749 | 2191 | 2751 | 2395 | 2961 | 2512 | 2831 | 3538 |
| 8plex3 | 2850 | 2025 | 2188 | 2317 | 2391 | 2852 | 3540 | 2194 |
| 8plex4 | 2106 | 2387 | 2867 | 2864 | 2401 | 2747 | 2746 | 2453 |
| 8plex5 | 2044 | 2484 | 2267 | 2755 | 2752 | 2741 | 2045 | 3532 |
| optimized set 3 | | | | | | | | |
| 8plex1 | 2194 | 2391 | 2191 | 2749 | 2745 | 3538 | 2861 | 2961 |
| 8plex2 | 2166 | 2188 | 2859 | 2212 | 2864 | 2746 | 2383 | 2752 |
| 8plex3 | 2484 | 2401 | 2850 | 2852 | 2512 | 2755 | 2106 | 2044 |
| 8plex4 | 2867 | 2453 | 3532 | 2025 | 2741 | 2267 | 2317 | 2387 |
| 8plex5 | 3511 | 3534 | 2751 | 2747 | 2395 | 3540 | 2831 | 2045 |

Without the use of said invention, the selection would have been performed randomly and tested for successful application later. Three randomly chosen subsets are shown here.

TABLE 4:

| | | | | | | | | |
|--------------|------|------|------|------|------|------|------|------|
| random set 1 | | | | | | | | |
| 8plex1 | 2191 | 2194 | 2267 | 2741 | 3534 | 3511 | 2749 | 2747 |
| 8plex2 | 2391 | 2484 | 2867 | 2852 | 2453 | 2512 | 2025 | 3538 |
| 8plex3 | 2746 | 2212 | 2755 | 2045 | 2044 | 2188 | 2961 | 2864 |
| 8plex4 | 2831 | 2383 | 3540 | 2859 | 2861 | 2395 | 2401 | 2317 |
| 8plex5 | 2106 | 2751 | 2387 | 2745 | 2752 | 3532 | 2850 | 2166 |
| random set 2 | | | | | | | | |
| 8plex1 | 2045 | 2106 | 2212 | 2745 | 2044 | 2749 | 2752 | 2391 |
| 8plex2 | 2025 | 2831 | 2401 | 3540 | 2395 | 2484 | 2453 | 2961 |
| 8plex3 | 2194 | 2859 | 2746 | 2512 | 2267 | 2864 | 2861 | 2751 |
| 8plex4 | 2383 | 2166 | 2747 | 2387 | 3532 | 2741 | 2867 | 2852 |
| 8plex5 | 3534 | 2755 | 2850 | 2317 | 2191 | 3538 | 3511 | 2188 |
| random set 3 | | | | | | | | |
| 8plex1 | 2484 | 2850 | 2741 | 2747 | 2755 | 2745 | 2025 | 2746 |
| 8plex2 | 2383 | 3534 | 2861 | 2751 | 2749 | 2391 | 2188 | 2191 |
| 8plex3 | 2194 | 3538 | 2512 | 2961 | 2864 | 2867 | 2831 | 3532 |
| 8plex4 | 3511 | 2045 | 2387 | 2212 | 2166 | 2267 | 3540 | 2401 |
| 8plex5 | 2395 | 2317 | 2859 | 2453 | 2852 | 2106 | 2752 | 2044 |

The sequences of all of those amplificates and the according primers are given in the sequence protocol (primers SEQ IDs 81-160; amplificates SEQ IDs 1-40). SEQ IDs refer to the internal numbers used in these tables as is shown in TABLES 1 and 2.

To show if the use of the design method described herein was superior to the common method of selecting primers for simultaneous amplification randomly said multiplex PCRs were performed. This example hereby demonstrates the advantage of the method which is subject of the invention:

A total of 40 amplicates (with lengths ranging from 187 - 499 bp) were partitioned into five 8-plex PCRs using either of two strategies.

5 First: the grouping was based on the invention using said "optimised sets" ("designed group").

Second: the grouping was done without using the selection criteria established by this invention using the "random
10 sets" ("control group").

Whether such grouping can improve the success rate of mPCRs was subsequently tested experimentally by comparing the number of true and false positives and false nega-
15 tives for each of the two classes.

Each of the five mPCRs (multiplex PCRs) contained 8 primer pairs specific for 8 amplicates with one primer of each pair being labeled with a Cy-5 fluorescent tag.
20 Only fragments that performed successfully in sPCR (singleplex PCR) using bisulfite-modified human DNA from whole blood were included in this study. Isomolar primer concentrations were used in a 20µl PCR reaction volume and cycling was done for 42 cycles using a 96-well micro-
25 titer plate thermocycler.

Group assignments for the "optimized" and "random" groups were done in triplicate and all mPCRs were run at the same time such as to minimize experimental variation in
30 PCR performance.

A mixture of the amplicates that were expected to be generated in a specific mPCR reaction but were generated in eight corresponding sPCR reactions was called sPCR-
35 pool. Electrophoresis of sPCR-pool amplicates and mPCR amplicates was done simultaneously using the ALFexpress

system (Amersham Pharmacia). In order to obtain the best comparability for mPCRs with their respective sPCR standard, these products were electrophoresed next to each other on the gels.

5

Figures 1 and 2 show examples of these results as electropherograms, given as ALFexpress output files.

Success or failure scoring for each mPCR was based on assessing the number of generated or absent fragments compared to their respective pool of sPCR fragments. Only fragments with peak areas equal or larger than 8% of the largest peak within one electropherogram were included into the analysis.

15

Figure 1 illustrates a result of an 8-plex PCR based on a primer combination from the "optimized set". The top graph in the figure shows peaks of size standards only. The second graph in the figure shows the electrophoresed mixture of the products from 8 singleplex PCRs. The third graph shows the products resulting from a multiplex PCR employing one of the optimized sets of primer combinations. By comparing these graphs it becomes visible that, in this specific example, there is only one false negative (FN) and three false positives (FP), whereas there are eight true positives (TP).

25

Figure 2, however, illustrates a result of an 8-plex PCR based on a primer combination from the "control set". The top graph in the figure shows peaks of size standards only. The second graph in the figure shows the electrophoresed mixture of the products from 8 singleplex PCRs. The third graph shows the products resulting from a multiplex PCR employing one of the randomly chosen sets, as is the state of the art. This graph clearly shows that, there are eight false negative and six false positive

30

35

peaks, whereas there is only one true positive. Hence, for this specific example we have demonstrated the superiority of the design method.

5 A more comprehensive view on the results is given in **Figures 3 and 4**.

By applying the Wilcoxon rank sum test for the determination of false positives or false negatives as follows, it
10 becomes evident that the optimized set resulted in a more reliable amplification experiment:

data: False negatives (FN)
p-value = 0.02602 rejection of null hypothesis
15 null hypothesis (H0): true if median of designed set equal or greater than of control set alternative hypothesis (H1): true if median of designed set less than of control set

20 data: False positives (FP)
p-value = 0.06711 rejection of null hypothesis
null hypothesis (H0): true if median of designed set equal or less than of control set
alternative hypothesis (H1): true if median of designed
25 set greater than of control set

data: True positives (TP)
p-value = 0.02146 rejection null hypothesis
null hypothesis (H0): true if median of designed set
30 equal or less than of control set
alternative hypothesis (H1): true if median of designed set greater than of control set

Figure 3 illustrates a summary of several such comparisons (as described in detail above). Six diagrams are
35 shown, that illustrate the numbers of false positives

(FP), false negatives (FN) and true positives (TP) for a number of 18 experiments. In the top row of figure 3 the results for experiments that employed the design method are shown whereas in the lower row results from experiments are shown, that did use the conventional method of random selection.

At the x-axis the occurrence of an event (like a false positive) per 8plex is given whereas the values of the y-axis indicate the frequency of an event like this occurring within the number of experiments performed.

For example, in the diagram title FN, a y-value of 0 indicates that the event did not occur in a single experiment, a y-value of four indicates that the according number of occurrences given as the x-value was found in four experiments (out of the 18 experiments considered for these analyses). The x-value indicates what kind of occurrence is counted; a x-value of three in this diagram indicates the occurrence of three false negatives. A data point with an x-value of 0 and an y-value of 9 means, that in the set of mPCR results considered, nine experiments showed 0 false negatives.

Figure 4 gives all of the data from the 18 multiplex PCR experiments of this example in one table. The letter A, heading the four columns presented on the left side, is indicating the results from multiplex PCRs of the designed group using the five optimized sets of primer pairs that have been designed and selected according to the invention. The letter C is indicating the results from multiplex PCRs of the control group using the five randomized sets of primer pairs.

The first column lists the identifying numbers of the experiments, the second column gives the numbers of true

positives (TP) within this experiment, the third column gives the numbers of false positives (FP) and the last column gives the numbers of false negatives (FN).

5 The average false negative rate (\emptyset FN) of the optimized group is significantly lower than in the control group. Complementary the average true positive rate (\emptyset TP) is significantly higher. The average false positive rates (\emptyset FP) of the two sets do not differ from each other sig-
10 nificantly.

This is due to the high deviation of false positives observed between individual ALFexpress analysis runs. Those 36 sets of amplicates have been analyzed on two separate gel runs. These runs were not designed to simply duplicate the results, but could be used to analyze whether
15 the average TP, FP and FN rates are similar, independent of the run, and the sets chosen. Only three of those sets have been duplicated, as indicated by the letters a and b for sets 11, 21 and 23. It turned out that the rate of
20 true positives as well as the rate of false negatives averaged over 18 sets per run were highly reproducible, 6.83 versus 7.33 and 1.44 versus 1.39 respectively. However, the rate of false positives was determined as 4.11
25 in the first run and 7.61 in the second run.

Taken together, it could be concluded that the overall success rate of amplifying 40 fragments within 5 groups of 8plex PCRs was significantly increased when the primer
30 grouping was based on the method being subject of this invention compared to an arbitrary primer grouping. The improved success rate of only 11% failures versus 24% in the random control group clearly becomes relevant when much larger numbers of mPCRs have to be established as is
35 the case in a high throughput laboratory.

Claims

1. A method for the amplification of nucleic acids comprising the following steps
- 1) isolating a nucleic acid sample,
 - 2) treating said sample in a manner that differentiates between methylated and un-methylated cytosine bases within said sample,
 - 3) amplifying at least one target sequence, within said treated nucleic acid, by means of enzymatic amplification and a set of primer molecules, wherein said primer molecules are characterized in that
 - a) each primer molecule sequence reaches a predefined measure of complexity,
 - b) every combination of any two primer molecules in the set has a melting temperature below a specified threshold temperature,
 - c) every combination of two primer molecules, under conditions allowing for one or more base mismatches per primer, does not lead to the amplification of an unwanted product when virtually tested using the treated and the untreated sample nucleic acids as template,
- and
- 4) detecting said amplified target nucleic acid.
2. A method according to claim 1 wherein said primer molecules do not contain nucleic acid sequences com-

plementary or identical to nucleic acid sequences of the target sequence which prior to treatment of step 2 contained a 5'-CG-3' site.

- 5 3. A method according to claims 1 and 2 wherein said set is comprised of at least one but not more than 32 primer pairs.
- 10 4. A method according to claims 1 and 2 wherein said set is comprised of at least one but not more than 16 primer pairs.
- 15 5. A method according to claims 1 to 4 wherein the primer molecules are reaching a specified value of linguistic complexity.
- 20 6. A method according to claims 1 to 4 wherein the primer molecules are reaching a specified value of Shannon entropy.
- 25 7. A method according to claims 1 to 6 wherein the nucleic acid sample is isolated from a bodily fluid, a cell culture, a tissue sample or a combination thereof.
- 30 8. A method according to claims 1 to 7 wherein the nucleic acid sample is comprised of plasmid DNA, BACs, YACs or genomic DNA.
- 35 9. A method according to claims 1 to 7 wherein the nucleic acid sample is comprised of human genomic DNA
10. A method according to claims 1 to 9 wherein said sample is treated by means of a solution of a bisulfite, hydrogen sulfite or disulfite.

11. A method according to claims 1 to 10 wherein said primer molecule comprises of at least one nucleotide within the last three nucleotides from the 3' end of the molecule wherein said nucleotide is complementary to a nucleotide of the target sequence that was converted to a different nucleotide by the treatment performed in step 2) of claim 1.
12. A method according to claims 1 to 10 wherein said primer molecule comprises of at least one nucleotide within the last three nucleotides from the 3' end of the molecule wherein said nucleotide is complementary to a nucleotide of the target sequence that was converted to a different nucleotide by bisulfite treatment.
13. A method according to claims 1 to 12 wherein each of said primer molecules is characterized in that the last at least 5 bases at the 3' end of said primer molecule are not complementary to the sequence of any other primer molecule in the set.
14. A method according to claims 1 to 13 wherein the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than 20% of the number of nucleotides of the primer molecule.
15. A method according to claims 1 to 13 wherein the number of nucleotides creating one gap, when aligning the primer molecule sequence with the template sequence, allowed for, when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than 20% of the number of nucleotides of the primer molecule.

16. A method according to claims 1 to 13 wherein the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than 10% of the number of nucleotides of the primer molecule.
17. A method according to claims 1 to 13 wherein the number of nucleotides creating one gap, when aligning the primer molecule sequence with the template sequence, allowed for, when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than 10% of the number of nucleotides of the primer molecule.
18. A method according to claims 1 to 13 wherein the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than 5% of the number of nucleotides of the primer molecule.
19. A method according to claims 1 to 13 wherein the number of nucleotides creating one gap, when aligning the primer molecule sequence with the template sequence, allowed for, when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than 5% of the number of nucleotides of the primer molecule.
20. A method according to claims 1 to 13 wherein the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is less than seven.
21. A method according to claim 20 wherein the number of mismatches allowed for is less than five.

22. A method according to claim 20 wherein the number of mismatches allowed for is less than three.
- 5 23. A method according to claim 20 wherein the number of mismatches allowed for is one.
- 10 24. A method according to claims 1 to 13 wherein the number of mismatches allowed for when virtually testing the amplification of unwanted products according to step 3 c) of claim 1 is determined by a pre-specified maximum melting temperature.
- 15 25. A method according to claims 1 to 24 wherein said primer molecules are used to amplify nucleic acid sequences that prior to treatment of step 2 comprised of more than eight 5'-CG-3' sites.
- 20 26. A method according to claims 1 to 24 wherein said primer molecules are used to amplify nucleic acid sequences that prior to treatment of step 2 comprised of more than six 5'-CG-3' sites.
- 25 27. A method according to claims 1 to 24 wherein said primer molecules are used to amplify nucleic acid sequences that prior to treatment of step 2 comprised of more than four 5'-CG-3' sites.
- 30 28. A method according to claims 1 to 24 wherein said primer molecules are used to amplify nucleic acid sequences that prior to treatment of step 2 comprised of more than two 5'-CG-3' sites.
- 35 29. A method according to claims 1 to 28 wherein the ability of said primer molecules to amplify an unwanted product is tested by means of electronic PCR.

30. A method according to claims 1 to 28 wherein the ability of said primer molecules to amplify an unwanted product is tested by means of electronic PCR, taking as template nucleic acid the coding strand of the treated sample, the non-coding strand of the treated sample and both of the strands of the untreated sample.
31. A method according to claims 1 to 28 wherein the ability of said primer molecules to amplify an unwanted product is tested by means of electronic PCR, taking as template nucleic acid the coding strand of the bisulfite converted human genome, the non-coding strand of the bisulfite converted human genome and both of the strands of the untreated human genome.
32. A method according to claims 1 to 31 wherein said primer molecules are used to amplify nucleic acids which are comprised of at least 50 bp but not more than 2000 bp.
33. A method according to claims 1 to 31 wherein said primer molecules are used to amplify nucleic acids which are comprised of at least 80 bp but not more than 1000 bp.
34. A method according to claims 1 to 33 wherein said primer molecules are comprised of 16 - 50 nucleotides.
35. A method according to claims 1 to 34 wherein said primer molecules do not form dimers with each other.
36. A method according to claims 1 to 35 wherein said primer molecules do not form loops or hairpin structures.

37. A method according to claims 1 to 36 wherein said primer molecules are complementary to target sequences which prior to the treatment performed in step 2) of claim 1 did not contain specified restriction enzyme recognition sites.
38. A method according to claims 1 to 37 wherein said primer molecules amplify regions of the treated nucleic acids which prior to the treatment performed in step 2) of claim 1 did not contain specified restriction enzyme recognition sites.
39. A method for designing primers according to claim 1, comprising the steps of
- a) selecting a pool of possible primer pairs per amplificate by means of a standard PCR primer design program using said nucleic acids as template
 - b) excluding those primer pairs which comprise of a primer that in combination with another primer molecule in the same set exceeds a threshold melting temperature
 - c) excluding those primer pairs which comprise of a primer that does not reach a specified level of complexity
 - d) excluding those primer pairs which comprise of a primer that in combination with another primer molecule in the same set, under conditions allowing for one or more base mismatches per primer, amplifies an unwanted product when virtually tested using the treated and the untreated sample nucleic acid as template.

40. A method for designing said primer molecules according to claim 1 and 43, adding the step of

5 e) excluding from the remaining confirmed primer pairs those pairs which in said amplification step do not result in the amplification of the intended product when performing a single PCR experiment.

10 41. A method for designing primers according to claims 39 and 40, wherein said template nucleic acids are masked for repeats and SNPs before designing said primer molecules and wherein said standard PCR primer design program considers one or more of the following
15 factors

length of amplificate, length of primer, melting temperature of the primers, dimer formation parameters, loop formation parameters, exclusion of unidentified
20 or ambiguous nucleotides in the primer sequence, exclusion of restriction enzyme recognition sites.

42. A method according to claims 39 to 41 wherein said
25 measure of complexity is a measure of linguistic complexity.

43. A method according to claims 39 to 41 wherein said
30 measure of complexity is a measure of Shannon entropy.

44. A method according to claims 39 to 43 wherein the
following step is carried out prior to performing
step d)

35 excluding from the remaining primer pairs those

pairs, which consist of a primer molecule that comprises of at least one CpG site.

- 5 45. A method according to claims 39 to 44 wherein the following step is carried out prior to performing step d)

10 excluding from the remaining primer pairs those pairs, which consist of a primer molecule that does not contain at least one nucleotide within the last three nucleotides from the 3' end of the molecule wherein said nucleotide is complementary to a nucleotide of the target sequence that was converted to a different nucleotide by the treatment performed in
15 step 2) of claim 1.

- 20 46. A method according to claims 39 to 45 wherein the following step is carried out prior to performing step d)

25 excluding from the remaining primer pairs those pairs, which consist of a primer molecule that contains more than 5 bases at its 3' end that are complementary to any other primer molecules' sequence in the set.

- 30 47. A method according to claim 39 to 46 wherein the following step is carried out prior to performing step d)

35 excluding from the remaining primer pairs those pairs, which amplify a nucleic acid that did not, prior to the treatment in step 2 of claim 1, contain at least two CpG sites

48. A method according to claim 39 to 47 wherein the following step is added before performing step d)

5 excluding from the remaining primer pairs those
pairs, which comprise of one primer molecule that in
combination with another primer molecule in the set
amplifies an unwanted product, when virtually testing
according to step 3 c) of claim 1 under conditions
10 allowing for a number of mismatching nucleotides of
20% of the number of nucleotides of the primer molecule.

49. A method according to claim 39 to 47 wherein the following step is added before performing step d)

15 excluding from the remaining primer pairs those
pairs, which comprise of one primer molecule that in
combination with another primer molecule in the set
amplifies an unwanted product, when virtually testing
according to step 3 c) of claim 1 under conditions
20 allowing for a number of nucleotides creating one
gap, when aligning the primer molecule sequence with
the template sequence, of up to 20% of the number of
nucleotides of the primer molecule.

- 25 50. A method according to claim 39 to 47 wherein the
following step is added before performing step d)

30 excluding from the remaining primer pairs those
pairs, which comprise of one primer molecule that in
combination with another primer molecule in the set
amplifies an unwanted product, when virtually testing
according to step 3 c) of claim 1 under conditions
allowing for four or less mismatching base pairs.

35

51. A method according to claim 39 to 47 wherein the following step is added before performing step d)

5 excluding from the remaining primer pairs those
 pairs, which comprise of one primer molecule that in
 combination with another primer molecule in the set
 amplifies an unwanted product, when virtually testing
 according to step 3 c) of claim 1 under conditions
10 allowing for two or less mismatching base pairs.

Sequence listing

<110> Epigenomics AG

5 <120> Method for amplification of nucleic acids of low complexity

<160> 160

<210> 1

10 <211> 322

<212> DNA

<213> Artificial Sequence

<220>

15 <223> 2025

<400> 1

20 aatcctccaa attctaaaaa cataaaaata acgcaaccca aaaacaaaaa acccctccgc 60
ccattaatta ctatacacta acgaaacttt cccgacccac aacgacgaaa ataaaaacaa 120
tcgctaacgc taaaaaacat caaaaacact acccaaccca aatatcgccg ccgcttccac 180
aaaactctac taaacgcgc cgccgccgct accaccgct ctaatecaaa ccacctccg 240
ccaaataaac cccgaaatcc taactcaaat atatatctct ccctccctct ccctccattc 300
gtcattttct cactcccttt cc 322

25

<210> 2

<211> 413

<212> DNA

<213> Artificial Sequence

30

<220>

<223> 2044

<400> 2

35

ggataggagt tgggattaag attttcggtt agtttcgtat tttttcgat tttttagtat 60

cgtttcgtat ttttcgtatt ttttttcggg ttattacggt ttttatgtga ttcgtttggg 120
taacgtcgaa tttagtcgcg tagcgttgta gtgaattttt tttttaaatt gtaataagtc 180
gttttttaag gtaattacgt tttttttggt ttttttttaa aaaataaaaa taaaaaattt 240
atagaaaaaa attcgcgagt ttagaaaaaa gaagtaattg gtagaagggt ttaattaagg 300
5 taaagagttg taaggcgaag ttaagaaaat gtaggtattt aaaaaatgta ggtaattttt 360
ataagggttt ttggggagag gtatatagag ggatttttgt gttgaaaaag att 413

<210> 3
<211> 347
10 <212> DNA
<213> Artificial Sequence

<220>
<223> 2045
15 <400> 3

aaccctttct tcaaattaca aaccttctta ccttcaaacc tcgactccaa caccaatccg 60
acaaaaaac ccaatctaataaaaatacgtt cccttctctac cattctctat tccattaacc 120
20 tatttcgtaa taaacgtaaa actaatcctc caaaattacc ttattaatta acttacatat 180
ttattatcta tctatccac caaaatacaa atttccgaaa aacaaaaatt taaaaaatc 240
tattttattc tatataattt tccatacca aacaccgtac ccgacacaaa ctaaaatccc 300
aatacacatc tcgaaacgaa aaaaccgtat ttccctaaaa cccaatc 347

25 <210> 4
<211> 283
<212> DNA
<213> Artificial Sequence

30 <220>
<223> 2106

<400> 4

35 ttgaaaataa gaaagggtga ggtagagagg ataatatagt tttagtttat tttttagtat 60
tttgtaatt ttttttaatt tttagttata aattcgagat ataacgtttt ttttttaaag 120

| | | |
|----|--|-----|
| | aggtcgcgtt ttttttgtgg tggtttttag ggattcgttt tagttttttt ttcgttttta | 180 |
| | gttttatata ttgggattat taggtattta agattttatt ttttaggtgg ttttttagc | 240 |
| | gtaggttgtt atttagtttt ttttaggga tttgggtag aag | 283 |
| 5 | <210> 5 | |
| | <211> 211 | |
| | <212> DNA | |
| | <213> Artificial Sequence | |
| 10 | <220> | |
| | <223> 2166 | |
| | <400> 5 | |
| 15 | tgtttgggat tgggtagggt tatcggggtt gggggggcgg ggtttgtggg taaggcgggc | 60 |
| | ggaggcgtgg atttttcgtt cgatgatagg gttggaggag gaaggggcgg gttgaagaag | 120 |
| | gggaagggtg gaagagtta gtcgggggta taaattgggt gaagcgttga ggttttagta | 180 |
| | ttttcgtttg aggagatagg taaaggttat g | 211 |
| 20 | <210> 6 | |
| | <211> 497 | |
| | <212> DNA | |
| | <213> Artificial Sequence | |
| 25 | <220> | |
| | <223> 2188 | |
| | <400> 6 | |
| 30 | ttttagattg aggttttagg gttaaaggat tttttttttt ttttagcgttg gttcgggaaa | 60 |
| | ggtaagtttc gggcgggagc gtacgtcgcg ttttcgaagt ttggtttttt cgttacgttt | 120 |
| | attttttgtt tttatttcgc gtttttttag gtttttttc ggtgaatcgg atgttttgtt | 180 |
| | agttttttat tttgcgtttt cggtcgcggt tcgggttttt cgtaaagtcg ttgttatttc | 240 |
| | ggagggttta gttagcgggt tttcggagggt tggtcgggta ggcgtggtgc gcgtaggag | 300 |
| 35 | ttgggcgcgt acggttatcg cgcgtggagg agatattgtt ttgtcgcgat gggggttcgg | 360 |
| | ggcgtttttt tacgtcgtag gtaagcgggg cggcgggttc ggtatttgtt tatcgggagt | 420 |

tttttttttt tttttttgtt gttgttggtt tgtatttagt tcgggggagg atagaagaaa 480
aaggaggtag aatggat 497

<210> 7
5 <211> 373
<212> DNA
<213> Artificial Sequence

<220>
10 <223> 2191

<400> 7

ggaggggaga gggttatgcg attttatttt tggttagggt cggggagggt tttgtttttc 60
15 gggagttttg ttcgggtttt ttggtcgtag ggttggtggg ttttaggtag gaacgagagg 120
gtgaggttta tatgtggttc ggcggtttag ggcggtttgt agcgttttta ttgtttcggg 180
tgtaggggt tgccgcgacg cggttagtta gtagcgagtt taggtcgcgt agattttatt 240
gatgagtttt gattttttagt atttttttta agttaagaag agtttagcgt atttttcggg 300
tgttttattt tagttttttt gtttttagttt ttttagttta ttttttttcg ttttggtttg 360
20 ggggtgtgtat agt 373

<210> 8
<211> 368
<212> DNA
25 <213> Artificial Sequence

<220>
<223> 2194

<400> 8

ttttgggaat gggttgtatc gagaggttcg attagtttta gggttttagt gaggggtag 60
tggaatttag cgagggattg agagttttat agtatgtacg agtttgatgt tagagaaaaa 120
gtcgggagat aaaggagtcg cgtgttatta aattgtcgtc gtagtcgtag ttatttaagt 180
35 gtcggatttg tgagtatttt gcgttttttag ttttcggata gaagttggag aatttttttg 240
gagaattttt cgagtttaga gacgagattt ttttaataatt attatttttt tttgcgtttt 300

ttattttgtcg ttcgttggga taaacgatag ttatagtttt ttgacgata ggatggaggt 360
taagggtta 368

<210> 9
5 <211> 352
<212> DNA
<213> Artificial Sequence

<220>
10 <223> 2212

<400> 9

15 ttgttgggag tttttaagtt ttgtgagaat tttgggagtt ggtgatgtta gattagttgg 60
gttatttgaa ggtagtagt tcgggtaggg tttatcgaaa gtttattcgt atatattagg 120
taatttaatt ttttattttg tgtgatagaa gtagtaggaa gtgagttgtt tagaggtagg 180
agggtttatt ttttgttaaa ggggggatta gaattttttt atgcgagttg tttgaggatt 240
gggatgtcga gaacgcgagc gattcgagta gggtttgttt gggatcgtc ggggtaggat 300
tcggaacgta ttcggaaggt ttttgtaag tatttatattg gaaggagaat tt 352

20 <210> 10
<211> 295
<212> DNA
<213> Artificial Sequence

25 <220>
<223> 2267

<400> 10

30 gtaatttgaa gaaagttgag gggaggcggg agatgttttg atttattagg gaaaacgtgg 60
acgttttttg ttgttatttt gtgaattgtg tgtatttagt tatttttgag taaatatttg 120
gagcgaggaa tttttgagtg gtgtgggagg gcggtgaggg gtagttgaaa gtcggttaaa 180
gttttcggag gggttggttt aggaaatatg attggtagtt acgagagagt taggggttg 240
35 acgtcgagga gagggagaag gttttcgggc ggagagaggt tttgtttagt tgttg 295

<210> 11

<211> 278

<212> DNA

<213> Artificial Sequence

5

<220>

<223> 2317

<400> 11

10

| | |
|--|-----|
| ggagttgtat tgttgggaga tttgggtgta gatgatgggg atgttaggat tattcgaatt | 60 |
| taaagttgaa cgttttaggta gaggagtgga gttttgggga attttgagtc ggtttaaagc | 120 |
| gtattttttt gtatatttat tcggtgttgg gcgtagggaa tttttgaaat aaaagatgta | 180 |
| taaagtattg aggtttgaga tttttggatt tcgaaatatt gagaatttat agttgtatat | 240 |
| tttagagttt atggtatttt agtgaaaatt ggggtttt | 278 |

15

<210> 12

<211> 285

<212> DNA

20

<213> Artificial Sequence

<220>

<223> 2383

25

<400> 12

30

| | |
|--|-----|
| tttgtatttag gttggaagtg gtcgttagtt tttcgtgtaa ttttattttt tggaaaagtg | 60 |
| gaattagttg gtattgttta gcgtgatttg tgaggttgag ttttaatagt ttaaagaagt | 120 |
| aaatgggatg ttatttttcgc ggggttcggt tttcgcgagg tgtttatttc gtatttgta | 180 |
| tgtaaaacga gggagcgtta ggaaggaatt cgttttgtaa agttattggt tttggttatt | 240 |
| agtttttatt taatgttttc gtgatgttgt tgttgattta tttgg | 285 |

<210> 13

<211> 380

35

<212> DNA

<213> Artificial Sequence

<220>

<223> 2387

5 <400> 13

| | | |
|----|--|-----|
| | gatttttga gaggaagtta agtggttttt tggttttttt cggatatttta tttaaggcga | 60 |
| | ttagttttaga attggttttc ggaagcgttc gggtaaagat tgcgaagaag aaaagatatt | 120 |
| | tggcggaaat ttgtgcgttt gggcggtgg aattcgggga ggagaggag ggattagata | 180 |
| 10 | ggagagtggg gattattttt ttgtttttta aattggggtg gttttttggg ttttcgattt | 240 |
| | ttttattttc gtgggtaaaa aattttgttt ttatcgggtt tacgtaattt ttttaagggg | 300 |
| | agaggaggga aaaatttgtg gggggtacga aaaggcggaa agaaatagtt atttcgttat | 360 |
| | atgggttttg ttttttagttt | 380 |

15 <210> 14

<211> 397

<212> DNA

<213> Artificial Sequence

20 <220>

<223> 2391

<400> 14

| | | |
|----|---|-----|
| 25 | tggggtagt ttaggatagg cgttcggggg acgcgtgttt ttattttacg gggacggtgg | 60 |
| | aggagagtta gcgagggttc gagggtagg tattttaacg aatggttttt ttggtgtttt | 120 |
| | ttgcgtttcg tcggtttatt ttttttttta taaaacgggt ttagttttta gtatttattt | 180 |
| | ttcgttatta attaggtatt tcgggagatt agttcgttcg aaagtttttg cgttatttcg | 240 |
| | cgggtttttt taggtggttt ttttagtttc gttttttttc gggatgtttg ttgattattt | 300 |
| 30 | cgagttcgcg tggcgtaaga gtacgagcgt cgagttcgtg cgcgttaagg ttgcgtgggc | 360 |
| | gggtatcgat ttttttgaga agtttttagtg ttttttaa | 397 |

<210> 15

<211> 547

35 <212> DNA

<213> Artificial Sequence

<220>

<223> 2395

5 <400> 15

tttttgtatt ggggtaggtt tcggtaggtg tatgggagga agtacggaga atttataagt 60
ttttcgattt tttagtttag acgttggttg gtttttttcg ttggagatcg cgtttttttt 120
aaatttttgt gagcgttgcg gaagtacgcg gggttcgggt cgttgagcgt tgtaagatag 180
10 gggagggagt cgggcgggag agggaggggc ggcgtcgggg cgggttttga tatagagtag 240
gcgtcgcggg tcgtagtata gtcggagatc gtagttcgga gttcgggtta gggtttattt 300
gttttcgtag cgtcggttcg cgtttttttg tcgtagttag cggtagagtgt cgcggttttg 360
agatttttcg gtcggatgcg cggcggtttt agtttttcgag cgtttgtttg tttcgttttg 420
ggttgttcgg gttttttggg tttttcggcg gttgtacgga gttaaggcgt ttcgttttcg 480
15 gcgttttttc cgggtgtcga tttaggttgt tcggagttcg gagtttatag aggagagaga 540
tagtttg 547

<210> 16

<211> 414

20 <212> DNA

<213> Artificial Sequence

<220>

<223> 2401

25

<400> 16

attagaagtg aaagtaatgg aatttcgatg taaatataat attatttttt tgtagagtta 60
ttttgagtat aataaatttg aattgtgtta atgttgggag aaaaaattta aaagaagaac 120
30 ggagcgaata gtagtttttt cgttcgttga tttagaaatag taggacgata ttttttcgat 180
tggaggagag cgtttgcgtt cgtatttagt tggcgttcgt ttttttggtt tttttttagt 240
cgtttttttt tttttttttc gcgttttagt tattcgggaa ggtattgcgg tagttgggtt 300
ttgattgggt gttttgaaag tttacgggtt attcgattgg tgaattcggg gtttttttagc 360
gcggtgagtt tgaaattggt cgtatttggt tttaaagttg gtttttgga attg 414

35

<210> 17

<211> 272

<212> DNA

<213> Artificial Sequence

5 <220>

<223> 2453

<400> 17

10 gggatggggtt attagttgta aatcgtggaa ttttttttga tataatgaaa agatgagggt 60
gtataagttt tttagtaggg tgatgatata aaaagttatc ggagtatttt ataaggtata 120
aattttttaga gatagtagag tatataagtt tttaggataa gagttaggaa gaaattatcg 180
gaagggaatta ttttattgtg tgtaaatatg atttttaagt tggtcgtggt ttttttggtg 240
gtttttttga tttttgtagt tttgtgtgaa gg 272

15

<210> 18

<211> 391

<212> DNA

<213> Artificial Sequence

20

<220>

<223> 2484

<400> 18

25

taattgaagg ggttaatagt ggaatttggt tgggtgtttg ttaaattttt ttttttggtt 60
ttgttttggtg tttttttttg aagggaatttt ttttcgtttt tgtaataaga ttttttataa 120
agtatagatt ttttatttta tttcgcggtg tttgtatcgg gttttattgg ttttaggagt 180
tgaatatttt tttaggtata tataggtggg atataaataa gggtttttga attattattt 240
30 ttttattacg atagtaattt aaaatgtttg ggaagatggt cgtgattttt ggagttttta 300
atatattttg gataatggtt gtagtttgta agttattttt ttttatttgt tttaaatggt 360
agtatttaat tttagttttg gttttggttt t 391

<210> 19

35 <211> 430

<212> DNA

<213> Artificial Sequence

<220>

<223> 2512

5

<400> 19

agtggttttg gagtttagat gtaatataat gattgatatt ggtatagtat atttattttg 60
tttttgtaaa taaaatggta tatgtgatgt ttttttttgt ttttttgat ataaaataat 120
10 atttgttttt atttattatg ttttatgtt tttattttgt atgttaggag ttaagtattt 180
tgtatgtatt aatttatttt gtttttataa taatttttat atgtaggaa tattatagtt 240
attttatgaa tgagtcgagg aaggattga gacgttaagt aatttgttta aggttacgta 300
gttagtaagt ggtagagtaa gaattattat ggttttataa gtttaggaaa aagtttgaaa 360
gaattaaaat gttaatagcg gggattttta ggaagtattg aagaggttat gggagaagtt 420
15 tttattttgt 430

<210> 20

<211> 475

<212> DNA

20

<213> Artificial Sequence

<220>

<223> 2741

25

<400> 20

laggggaaaa gttagagttg agaggttggg gcgcgacgag tttggatata ggcggggat 60
ttaagttttt ttcgttttagt taataattgt gtttttttta ggaaggcgtg aggaaatggt 120
ttaattaatt tttgtatttt ttttttgaa tttgggttgt atttttttat ttattgtaaa 180
30 ttttataatt tatttagggg tttttttagt gtttgttttt agcggtttcg gtgtttattt 240
attagtgttg tttttttttt ttcgtaagat tgcgttttag ttttagtttt ttttttcgcg 300
ggtgtttttt aaatcgtttt attattttcg ggttttagga ggcggaatcg tgtttgtttt 360
tcggtttttt taagaggcgt cggttttatt ttttttagag tcgcggtttg acgcgagatg 420
atagtaacga gttcggtagt tttatgtaaa taagcgtttt tttgtgggtt aatgg 475

35

<210> 21

<211> 412

<212> DNA

<213> Artificial Sequence

5 <220>

<223> 2745

<400> 21

| | | |
|----|---|-----|
| 10 | at tt t t a g t t t g t g a a a t g g g a t t t a g g a t t t a g g t a g a g g t g c g t t t t c g g t t t g g g g a t | 60 |
| | c g a g t a t t t t g t g c g t t t c g g t a a c g t a g g a a g a t a g c g t t a t t g a t a t t t t a g a g a t t a | 120 |
| | g c g g g t a t c g t t t g g a g g c g t t t t a t t a t t t g g c g g t t t c g g g t t c g c g t t t a t c g c g | 180 |
| | t t a t a a g a t t t a c g t t c g a a t t a c g t g a t t a g g g t c g t g g t t t c g t t t c g t t t t c g c g t c | 240 |
| | g c g c g t c g t t t t c g g t a g g g g c g g a a a g c g g a a g t g t g g g a g g g t t t g c g g g c g g g t t t | 300 |
| 15 | a g g a g g t t c g c g g g a g g a t g g a g t a g t g a g c g g g t t t g g g c g g t t g t t g g t a g c g t t a t g | 360 |
| | g a g a c g g t a t a g t t g a g g a a t t c g t c g c g t c g g t g a g g g g t t a t t g g t t a a g | 412 |

<210> 22

<211> 484

20 <212> DNA

<213> Artificial Sequence

<220>

<223> 2746

25

<400> 22

| | | |
|----|---|-----|
| | g t g g g t t t t g g t a g t t a t a g a a g t t a t c g c g t t g g c g g g g a g g a g g g g g a t c g a t g c g g | 60 |
| | t t t a t g t t t c g g t a g t t t t a t t t t t t t t g t t t g c a a g g g t t t t t g t t c g g c g g g a g g a | 120 |
| 30 | g a g a g g c g c g t t t t a t t c g g g t t t t t t a t a t t t g t c g t c g t t t g g g t c g a t t t c g c g g g | 180 |
| | t t t c g t t c g g c g t t t t a g t c g a t t t t c g t t t a g t t t t c g g g t t t a t g g g c g c g g t t a g t a g | 240 |
| | g g c g g g t t a g g g c g g c g g g g c g c g a t a t t g g a g g a a g t g c g g t c g t t t g t t c g g g c g c | 300 |
| | g t t a a g g a a g t t g t t t a a a a t g a g g a a g a g t c g c g g g t t c g g c g g t t g a g g t a t t t c g g | 360 |
| | c g g c g g t t g g a g a c g a g g a g g a g c g g g t g t t t t c g c g t t g c g t t c g t t t t c g t t t t a t t | 420 |
| 35 | t g g c g t a g g t a g g t g t g g t c g c g t t t t t a t t c g g t c g g g a t t t t t t g g t a a g g a g a g g a | 480 |
| | g g t t | 484 |

<210> 23

<211> 476

<212> DNA

5 <213> Artificial Sequence

<220>

<223> 2747

10 <400> 23

taggatgggg agagtaatgt tttcgagtag aatagggtgg ggtttttaga ttatTTTTTT 60
TTTTTTatag ttggTTTTat tttatcgatt ttattaaagt ttttttgga gtatTTTtaga 120
gaagagttac gtttaggtcg ggTTTTggtt gtttggttta cggcggaatt ttagtatta 180
15 cgtttcgtac gtcgggttta aagtatgttt agtgaaggag taggtattta ttgtagatg 240
gagttatttt tttagatttg gggTTTTTTT ataacgatgg ttatgtttgg tatggaagtt 300
TTTTtagaag ttaatagtag gaaataaggg ttaatagtat ttaattgtgg agtaaggttt 360
aaattttagt tttgtatttt aatcgtttcg aatttgTTTT tttattgtag aggcgaaaag 420
gttaatatta tttattttcg gagggttatc gtggagaatg gaagttggat aagttg 476

20

<210> 24

<211> 419

<212> DNA

<213> Artificial Sequence

25

<220>

<223> 2749

<400> 24

30

tcccacaaaa actaaacaat tattacaaat tcaaaaaacc ccgaccaatt tttcaaaaat 60
ttctcctcct cttttccccc taaaactcgt aatactttta ctctactttc aaaatacatt 120
aaatctccta ctttataact actttaaaac caacaaatac tctaatatat ataattcaaa 180
ttatacaaat ttcacgaata aatttaactt tattttttta attaattaaa aaacaaataa 240
35 tattttaaaa aatattaact tataattatt tcaccctttt tacttttaac atttttatta 300
cttctcgacc ttttaactaa aatcaaatat atacttttaa cattttttta aataaaaaata 360

tccttttaat ttaataaaaa aacaaaattc tacataaaaa aaccccttca tctaaaacc 419

<210> 25
<211> 479
5 <212> DNA
<213> Artificial Sequence

<220>
<223> 2751
10 <400> 25

tttggagggt ttagtagaag ttatttttagg ggagggttcg ataggaagga aggtagggtt 60
gtcggagggg tatataggag tttttttttt cgttatagtg tttagggtta attgttttag 120
15 ttttttaggtt gggttaatag gatgggatag tttaggcgga aggaaatttg tggggaggga 180
tatttcgtag atagaagtag ggatatgggg tggggagagg taggaagagt tgcggggttg 240
ttgagttggc gtttttttag tagatttagg aggggcggtg ataggaggtt attttttttt 300
tattttcgta gttttgggtt tttttggtt tggttaatag tattattatt attattattg 360
ttgttgttcg ttagtttggg ttttagatat attagaaaa aattatcgga agatacgtat 420
20 agtattggta gtttttaaaa gaattaattt ttttttgtg tttattttgt gattattgg 479

<210> 26
<211> 484
<212> DNA
25 <213> Artificial Sequence

<220>
<223> 2752

<400> 26
30

atacaacctc aaatcctatc caaaccacca aaacatcaca ctcgaaactt attctacata 60
tttttacttt tacctccac taatactaatt tcttccgtaa aacaacctaa atcccttcaa 120
atacttaata ttttttctca aatactacca taaaaccaa tctccaccgt cttaaaacat 180
35 tcctttttta aaataaaaaa tatatatcgc tccttttata taatttacat tctatcttaa 240
ataatttaac catcaccgta attcattcaa atctatttaa atcctacca tctcaacttc 300

aatccatttc attcttttaa atctaatacg caattacctc caacaacttc atcacaaatc 360
actcacaaaa ataaccttaa tcctaaaatt tatttacgaa aaacacactt actaaatata 420
taacaaatat acaaaaaaca caaaataaaa caacaaatct aaaaacaaat aacttccttc 480
tccc 484

5

<210> 27
<211> 371
<212> DNA
<213> Artificial Sequence

10

<220>
<223> 2755

<400> 27

15

ggaagatgag gaagttgatt agatattaag gatgagcgga tgatttaata ggtttttttg 60
ttaagatttg gttgggtagg tgaaagataa agtcgaggag tggttatggt gtggtataga 120
agaagggtta gaggacggtt ttgtttattt ttttatgttt gagttttttt ttttgtgaaa 180
tggggataat aagagtcggt atatagggaa ttgttgtag gattaaatga gataatgtat 240
gtgaaacggt ttggtttagt gttttttagt aaatgggtac gatttgcgga gtggggattt 300
gaatttacgt ttggcgggat gtttaagttg ttattttgat cgtagggag ttttagagga 360
tagggttgta g 371

20

<210> 28
<211> 186
<212> DNA
<213> Artificial Sequence

25

<220>
<223> 2831

30

<400> 28

35

ttagtagggg tgtgagtgtt ttgattagaa ttattttttt ttgttagaat ttgatgtaat 60
tcgaatgttt ttatttttgt ttgaagggtt taaataataa attaggtttt gtcgtgttat 120
tatgggggtg gttatatttt gtatttagga aataggtacg gtagggttga gatagaagtt 180

ttgttt 186

<210> 29
<211> 300
5 <212> DNA
<213> Artificial Sequence

<220>
<223> 2850
10 <400> 29

ttatagggtt gagtttgga tcgaggtgag agtcgtcggg ttgggagtga gggagatggg 60
aataaggctc tcggtggcg agggagtcg aggaattcg gggattggg aggtttggg 120
15 cggcgcggtt tggtcgggtt gggatcgggt ttctcggtta gacgttcgcg atgttggtat 180
tttttggtat tttttatttg ggttttaggg gttcgttttt gggtagtttg gagtttttcg 240
aggtgggagg atcgggcgga ggtggaggaa gttttttttt ggaagatttg ttgtttggtt 300

<210> 30
20 <211> 321
<212> DNA
<213> Artificial Sequence

<220>
25 <223> 2852

<400> 30

tgaaaatgaa ggtatggagt ttggtgttaa aagaaatfff ttttaaaaat taaataataa 60
30 tattagagta aagtttttag ggcgagataa ggagtgttaa taaaataagc ggaaattcga 120
gaagcgttaa tgttttaaag ggttaatgat tatatataat ttacgtagtt aacgtgttaa 180
aatatatata cgtatttttt ttttttaaata aaagtaggaa agcggatttt gtatgagggg 240
cgggttgctc atttagtagt ttttttcgga tagttcgttt tgattttttt tggttggctc 300
tggagggatt atatggtttt a 321

35 <210> 31

<211> 398

<212> DNA

<213> Artificial Sequence

5 <220>

<223> 2859

<400> 31

| | | | | | | | |
|----|------------|------------|------------|-------------|------------|-------------|-----|
| 10 | tatgttttgg | tttgttttga | gatagagttt | cgttttgtcg | tttaggttgg | ttaaaagata | 60 |
| | gggttttagt | cgggtgcggt | ggtttacgtt | tgtaatttta | gtattttggg | aggtcgaggc | 120 |
| | gggcggatta | tttgaggttc | ggagttcgag | attagtttgg | gttaatatgg | cgaaacgttg | 180 |
| | tttttattaa | aaataataaa | aattatttag | gcgtgggtggc | gcgtatttgt | aatttttagtt | 240 |
| | attcgggagg | ttgaggtagg | agaattat | gaatttagga | ggtagacgtt | gtagtgagtc | 300 |
| 15 | gagatcgcg | tattgtat | tagtttggc | gatagaggga | gatttcgtt | taaaaaagg | 360 |
| | aaaaaaaaa | aaaagaaa | aaataaa | agt | gatgggg | | 398 |

<210> 32

<211> 347

20 <212> DNA

<213> Artificial Sequence

<220>

<223> 2861

25

<400> 32

| | | | | | | | |
|----|------------|------------|------------|------------|------------|------------|-----|
| | gggtgtagaa | gtgttttagt | tttttttcgt | tggggttggg | agtttgggta | ggttagtttt | 60 |
| | atttttttta | agttcgtttt | tggtttttcg | gtttagtttc | ggttattatg | tttcgtaga | 120 |
| 30 | ttatttttgt | gggttttagt | tgtttggtat | tgtggaggga | aaagaatgat | cggttcgttc | 180 |
| | gataggttta | ggtaatacgg | ttgttggtat | tttcggtttg | tagttttaag | atttttgaaa | 240 |
| | gcgggtttgt | agtggattta | ttttaataga | tggggaggga | ttgagtttga | ttaaagagtt | 300 |
| | agaaatgatt | ggagaatgta | ttttttgtta | ttgttgtaag | gggagaa | | 347 |

35 <210> 33

<211> 291

<212> DNA

<213> Artificial Sequence

<220>

5 <223> 2864

<400> 33

10 tcccccttcca actatatctc tcacccaaaa ataacttcta actctcgtat tcatctaaaa 60
ctcctccttc catataccaa caattaacta taaccctcc aaaaacgctc catctccaaa 120
tatactccca catccaaacc acgaaccctt caccgatca catacttcat acacctataa 180
ctcgcactc cccaaatata cctctaactg acaactatta ccccttcccc cgattataac 240
cctataactc gccacatata actataacta aaacttcctt aaaacactct c 291

15 <210> 34

<211> 389

<212> DNA

<213> Artificial Sequence

20 <220>

<223> 2867

<400> 34

25 aaaacaaaa cataaaccaa aaaccaaact cgaaccgaaa acaataaccg caacgcccga 60
aaactaaacc cagcagcgc taacaacgcg aaccgaacta cgaaaacgat caggtcaacg 120
tccgttcaa accgactaac aatctccgtt ctacattaac gtcaacactc ccgttaaaaa 180
taatacatct ctcccatacc aaaaaaactt aaatactact aaaaaccaac cctccgaata 240
ctaccaaacc gagctcacc cgccaccttc atcttccctt ctcctttacc ccaaaacaac 300
30 cgaaaatata taattaaatt cccctaccc ataaaaaac caaaaataaa aaactaacga 360
cctactcgat ctcaacaaac cctcctaatt 389

<210> 35

<211> 272

35 <212> DNA

<213> Artificial Sequence

<220>

<223> 2961

5 <400> 35

aatggttgat gatttttggt ttttttcgtc gtcggagagc ggtgtttcgg aggcggcgga 60
ggaggattcg gcggtcgttt ttttggttta gtaggagagc gagattgtag gtatagagaa 120
cgacgagggt ttcggggtat ttgtcggtag ttatgcggtt ttcgcgtagt cgggttttac 180
10 gagtgggggt gagttagcgc ggggtttgga gaggggttta gggcgcgtag tcgggggatt 240
tcggtcgggg tttaggggta tagggaagag ag 272

<210> 36

<211> 371

15 <212> DNA

<213> Artificial Sequence

<220>

<223> 3511

20

<400> 36

agttagaaga ggagttagga tgggtttcgg gtagtttaaat agtatagttg aagttttaat 60
tattatgtta atagtttttt ggttttatat attttatggg aagaggaaaa taaaaggta 120
25 tttatttgta ttttttttta ttttgatat aagaagtaga atttttttta tatgatttat 180
gtttatttaa tacgttattt tgaaatttat taataaaatt ttttaagcgt tagaaaattg 240
ttagtggttt tttttatttt tttttatttt tttttgtgtt attaatattg tttttttttt 300
ttagaagggt gtcggaatag taaatattta ttgatatgtt ataattattg gaaaatgggt 360
attggaaaat t 371

30

<210> 37

<211> 457

<212> DNA

<213> Artificial Sequence

35

<220>

<223> 3532

<400> 37

5 tgtttagtaga gtttttaggga gggttttattt tttatttttta tttaaagtgtt tatttggttg 60
 gggtgggggtt ttgttttgaa ggggaagggtt taagggtgtt tttagcgtgt ttttttattt 120
 tgattgtttt tggcggggcg ggggtgtttt tgttatttag ttgtataacg gtttaggaagg 180
 gtttaaatta ttttttagggt taatttaagg tcgttttttg ggtttgtata tttttgtgtt 240
 gagtgcggat cgggagagggt tgttgaagat aggaggggat aaatggggga cgaaggggtt 300
10 cgagggaggg gattgaagga tttgggttaa gtcgggagtt ttcgagggcg gaggtaaaac 360
 gtatttgat tttgttagtt ttaaattttg tttttattgt tgtaagtttt ttagatcgag 420
 gattttcggg ttgaggggtg ggtaaggata ggtagtg 457

<210> 38

15 <211> 476

<212> DNA

<213> Artificial Sequence

<220>

20 <223> 3534

<400> 38

25 tttttgtttt tatggggtgt atatttaagt agttgaaata gatagtgaat aaataaaaaa 60
 ggataataat tttaaataat aatgatgtta tcgggttaggt gtggtggttt atgtttataa 120
 tttttagtatt ttgggaagtt aagttaagcg gattatttga ggtaggaggt ttaagaatag 180
 tttggttagt atggtgaaat tttattttta ttaaaaatat aaaaattagt tagatatggt 240
 ggtatatatt tgtaatttta gttatttggg aggttgacgt aggagaattg tttgagttcg 300
 ggaggtggag gttgtagtga gttaagattt gataggtttt tagtattatt gtattttaga 360
30 ttggttgata gagcgagatt ttgttaaaaa aaaaaaagtt ataaatagat ttaatatagg 420
 taatatgata gggagggagg gataggggag taggggtggtt aaggaaggga tatttta 476

<210> 39

<211> 458

35 <212> DNA

<213> Artificial Sequence

<220>

<223> 3538

5 <400> 39

tgggtagtat ttttgttgg ttttttttat attataaggt tacgtagagt tggcggaggg 60
ttatggttttt atttatgtta ggtgttttta atttggttaag gaaatgtaat ttacgtgaat 120
tttaataggt agtgaagtat cgtttttttt tgatttttagg tagggtgaag aaaatgggat 180
10 agtagtacgg ggtgcgggta taaacgtata attttgtttt tttagacgta gaggttgtggg 240
gttgtgagaa tgtaggagg aggtaagaaa gggcggtttt atgggggggtt tgtagggtgg 300
gataagttta agaggttttt atatttaggt ttggtggggg aggtgagttt ttggtttatc 360
gagggggttt ttttttgttt tcggaatat tgtagttttt atttttatcg ttttttcgtt 420
gcggggattt aggggcgtga ggatgagaga gtttttag 458

15

<210> 40

<211> 405

<212> DNA

<213> Artificial Sequence

20

<220>

<223> 3540

<400> 40

25

agtggtttag gagtatttgg ttatttttcgg gaaaaatcgg tttggtaaag gttttttcga 60
gggtacgcgt ttttcggata gtgaggtagg atttaaattt tttcgttaat attatatttt 120
tcgtattttt gtagtgtttg tatttttagg ttttattatt ttttcgtatt ttttagggag 180
aagttttcga cgtttttattt tttttggaag ggtgttgttt ttagagattt ttaggttaat 240
30 ggtttaattt tagtgttttt aggggagagg ggggtgtaga aaaatagttt gggttataaa 300
agagggtcga gggttgtgag atttcggagg tatcgacggg aagcgagacg gagaatagga 360
gggtaggacg gggttgagggt gggggatatt gtagatggag ggagt 405

<210> 41

35 <211> 2501

<212> DNA

<213> Homo Sapiens

<400> 41

5 ccagttccag tcccgggtcc tgtggccgcc ctgccggcga cctgcgag agcgagtctt 60
agatacccag tcccagccc cgagttgtta ttccctcgct gtagttaaga aggaggagat 120
caattaaggg catcttagaa gttaggcggt cccgctgcct cctttgagca cggaggccac 180
caaccccta gggggaagag atgtagcgcg aggcaggggt gtcgtgctaa gaaatttcga 240
cgcttctggg gactgaggac aaaggtgcgg acacgacccc ggggtacctg gagtccgtg 300
10 actcgcgcca cggacggcac acctaggggc taatttctgc tctgcctcaa agaacctcaa 360
gctagagtcc ttgcctccgc ccacagcccc gggatgccgc tgcgctgc accgcacagg 420
cagcgcccgg accggctgca gcagatcgcg cgctgcgcgt tccaccggga gatggtggag 480
acgctgaaaa gcttctttct tgccactctg gacgctgtgg gcggcaagcg ccttagtccc 540
tacctctgct gagctgaacg ctcaggcaca gtggaactga aaccgggttc tgcgggatgt 600
15 gagagctgtt gaggtcacgc gtaattgggt gtgatggagg gcgcctgttc gtgatgtgtg 660
caggtttgat gcaagcaggt catcgctgtg cgagtgtgtg gatgcgaccg cccgagagac 720
tcggaggcag gcttgggaca cgtttgagtg aacacctcag gatactcttc tggccagtat 780
ctgtttttta gtgtctgtga ttcagagtgg gcacatgttg ggagacagta atgggttttg 840
gtgtgtgtaa atgagtgtga ccggaagcga gtgtgagctt gatctaggca gggaccacac 900
20 agcactgtca cacctgcctg ctcttttagta gaggactgaa gtgcgggggt gggggtacgg 960
ggccggaata gaatgtctct gggacatctt ggcaaacagc agccggaagc aaaggggag 1020
ctgtgcaaac ggctcaggca ggtgatggat ggcagggtag gaagggggag gtccagaggt 1080
ctggatggag gcttccgcat ctgtaccttg caactacccc ctcaggccca gcaggtcatc 1140
ggccccctcc tcacacatgt aatggatctg aagagtaccc cgggacagtc cggggagatg 1200
25 gagattcgga aagtatccat ggagatctta cagaatcccc tgtgcggacc aggaaactct 1260
tgtagatccc tgccatctg aggccaggc gctgggctgt ttctcacaat attccttcaa 1320
gatgagattg tgggtcccat ttcaaagatg agtacactga gcctctgtga agttacttgc 1380
ccatgatcac acaaccagga attgggcaa ctgtaattga actcctgtct acaaaagttc 1440
ttgctcccag ctccgtctct tgtttccac gagccctggc cctctgtggg taataccagc 1500
30 tactggagtc agatttcttg gggccagaac ccacccttag gggcattaac ctttaaaatc 1560
tcacttgggc aggggtcttg gatcagagtt ggaagagtcc ctacaatcct ggacctttc 1620
cgccaaatcg tgaaaccagg ggtggagtgg ggcgagggtt caaaaccagg ccgactgag 1680
aggtgaaatt caccatgacg tcaaactgcc ctcaaattcc cgctcacttt aaggcggtta 1740
ctgttgggtg cccccacat cccccacat ttccatcaat gacctcaatg caaatacaag 1800
35 tgggacggtc ctgctggatc ctccaggttc tggaagcatg aggggtgacg aaccagggg 1860
caaaggaccc ctccgcccac tgggtgctgt gcactggcgg aactttcccg acccacagcg 1920

gcgggaataa gagcagtcgc tggcgctggg aggcacacaga gacactgcc agcccaagtg 1980
 tcgcgcgcgc ttccacaggg ctctgctgga cgccgcgcgc gccgctgcca ccgcctctga 2040
 tccaagccac ctcccgccag gtgagcccg agatcctggc tcaggatat gtctctccct 2100
 ccctctccct ccattcgtea tttctcact cccttctc ctctccctct ctctccgta 2160
 5 gtctcttcat cagatagtct ctgttagtcc gcgatttata ccaggctcgt gccctagggt 2220
 ggatcggaaca gtctcaatcc ccggctcgc tcttctgct cggctgcgga ctccagtctt 2280
 actctctcgc actgcacaca ggcttaggcc agtctcgga cactcaggct cccagggac 2340
 cgcgcacaga gcctgaggca agagaaactt tccgcagac gtgcgatcag ggacggcgtc 2400
 tggagcccag cagtcccagg gaaattgggt cagaacctgg aacagagcgg atgggtggca 2460
 10 aataggcacg acgactgagg gacaagcagc cctaaactgc a 2501

<210> 42

<211> 2501

<212> DNA

15 <213> Homo Sapiens

<400> 42

agatttactc aaatttaaga atgagaatac aaatccacat cttgaagtgt ttcacagaaa 60
 20 ggtctatctt aatgtctgga gtatatattt caatgaacat tcattttatt ttatttctct 120
 ccattctcga atcaagcaat cttgaatcta aagtgctat gattagcact gaaaagacca 180
 ctggactatt aattgtgtga ctttgggaca gtaactttct gcaccttagt ttgtttacat 240
 gttatacatg aagggttgaag tctgattctg ctctgtgact atcattctaa acatctgatg 300
 aaatcaaatt tcagtgttgg gaatggtagt acaataaatt tactaagaat aaataattca 360
 25 ctgcaaaaac acattgattt ccaaatgatg taactgacag ttatattact gcagagggct 420
 gataaataac aaaagaaatg aaagatgcac atggtgagaa ctgaaattat cctgacaagt 480
 ctctacctg tttatcactt aaaatcaatg accatgctga atgcctacaa attacaaaat 540
 ataaaagaaa tcttataaat gcgcatgtac aggagtctaa gttactaaaa gttttaaagc 600
 ataaagttta accaaactaa tcaaagaagt tgagaggaaa aattggcttt catctttaat 660
 30 cactactgtt ttgaggctct atgtttaata taattttcta agtagaggct tcagagagaa 720
 gagttgtgag gatactttca tatttgtgta gaaggaaaag ttgcatcc attctagtat 780
 ccctagtgtt atactgatgt gcaccttga tttattttgt tcctattgta taaactcata 840
 cttgacttca aagaaaagga aaatccaaag tccctctttt ctaaggggac agaaatcctt 900
 tgtgtcaact gtttgacctt tttctctgta aggtcctatt ggaaatcttt tgtaacacaa 960
 35 tgcaggggac tcttccatgt gttgatgctg tttacacagt ggggtgggac tgactgaaga 1020
 aaaaaaatcg catatacgca tgaaagatta tggctcttatt tccggaaaagc atgaaagggtg 1080

attgatactt ccaagaagtc cctgttactc aggaaaatta tcaaattatc tactcagaga 1140
tacttggaag gactgaagga aaggaagaac gaagaaagca gaatctagac ttatgtggg 1200
agagatttgt ggcagaggaa aagtattctc tttgaatccg acaagggatt tgcctgggg 1260
aatttcctgt ccagcctttt attaccaggg tcttttgaag ccgggctccc cattgggcag 1320
5 ttccttgga gtgcagtggg gaattcttac actttccctc taggtccccg aaggatctcg 1380
ttttctcagt gtctctttca ggttggcagg agccttgagc ctgacacttc cctttgatgg 1440
gacaggcaag ctctgtgggc gcgtaaacac gctgtaacca agttctttgc tgattttaca 1500
gttttgtgtg ctcccagaaa gaagtgatcg tactcaattg tctattgctg gcctgcccc 1560
taagagcctg ggggctcctt tcccctaacc cagaactagc tgcacggggg gcggggaaat 1620
10 gggggtggg aaggagtggg agggcagtgg tttccgcgag cagagcgatg ttactgagtg 1680
agtccctgaa tggggagcgc tgctgtcccc aagccgattg gtacttcttg tcaggaagaa 1740
acgccaaagag gtgggagtgc ctggggaggg aggcaggcgg tccctaccgc aggcgcgggg 1800
agctgccttt ccgcccctcc gctgtcttcc caagcctgga ctcttaggag tggctgaagc 1860
tgcgagcgc ttttgagacc tgtgaatgaa cctcctcctc ctccctcctc cttcttctcg 1920
15 ctgagtctcc tctcggctc tgacggtaca gtgatataat gatgatgggt gtcacaaccc 1980
gcatttgaac ttgcaggcga gctgccccga gcctttcttg ggaagaactc caggcgtgcg 2040
gacgcaacag ccgagaacat taggtgttgt ggacaggagc tgggaccaag atcttcggcc 2100
agccccgcat cctccgcac cttccagcac cgtccgcac cctccgcac cttccccggg 2160
ccaccacgct tcctatgtga cccgcctggg caacgccgaa ccagtcgcg cagcgtgca 2220
20 gtgaattttc cccccaaact gcaataagcc gccttccaag gtaatcacgt ttcttttgtt 2280
cccccttaa aaaacaaaa caaaaaactt atagaaaaaa acccgcgagc ttagaaaaaa 2340
gaagcaattg gtagaaggct ttaattaagg caaagagctg taaggcgaag ttaagaaaat 2400
gtaggcactt aaaaaatgca ggtaactttc ataagggtt ttggggagag gcatacagag 2460
ggaccttggg gttgaaaaag attcagacaa aagaaccca g 2501
25
<210> 43
<211> 2501
<212> DNA
<213> Homo Sapiens
30
<400> 43

tgtgggtcat taatgcaatg ttatttaaga ctaggatttg gctgggcgca gtggctcacg 60
cctgtaatcc cagcactgtg ggaggccgag ccgggaggat cacctgaggt caggagttca 120
35 agaccagcct gaccaacatg gtgaaaccac gtcttacta aaaatacaaa attagccggg 180
catagtcaca tgcctgtaat ccagctact gggtagcctg aggcaggaga atcgcttgaa 240

| | | |
|----|--|------|
| | cccgaggagc ggaggcggag tttgcagtga gccaaagattt cacaactgca ctccagtctg | 300 |
| | ggccacaaga gcgaaaaccc gtctcaaaaa aaaaaaaaag actaggattt gacataaggc | 360 |
| | ctgaggggta ttcttttgtt ttgttttgcc ttgttttcaa gaggccaaaa tcttcacagt | 420 |
| | tgaaaatttc tgttgaacca cagagatttg aaccaactca gtttagaaag cctggggatt | 480 |
| 5 | tgaacaacgg tatggatcgg aaatctcttc atctgtcagt tttcatcatt ctaggcagta | 540 |
| | aaatagattt cccttttagga gcttttcacc gtttggggtt ctccagcagt gggatgtggg | 600 |
| | gaatcaaccc ttcttcgtct ccacccaaac attaggtggg agcaaggggt ggaagtaga | 660 |
| | gaaagtggat agaggtctcc agtggatatg ggatctttgt gtagaccagc acagtcctca | 720 |
| | gaaatctcat gcaagcaaca taggtactgt tatattttct agtggccacc ttttaaaaag | 780 |
| 10 | taaacaggtg aggcggggcg cggtcgtcac gcctgtaatc ccagcacttt gggaggccca | 840 |
| | ggcggggcga tcacgaggtc aagagatgga gaccatctg gtcgacacgg tgaaaccccg | 900 |
| | tctctactaa aaatacaaaa attagctggg catggtgacg cgcgactgta gtcctagcta | 960 |
| | ctggggaggc cgaggcagga gaatcacttg aaccctggag gtggagggtg ccacgctcca | 1020 |
| | ctacactcca gcctggcgac agagtggagac tccgtctcaa aaaaaagaaa gtaaacaggt | 1080 |
| 15 | gaaattaatt ttaataatat attttgttta acccaacgta tccaaaatac tatcatttga | 1140 |
| | aagtgtaatg aatataaaaa tattcatgag atatttttca ttctcatatc catactgtct | 1200 |
| | tggactctaa tgtgtatttt acacttacag cacaattaat ttgggactag ctacatttca | 1260 |
| | gctcaacaat agccaatagc atatgggata gcgcaataa actctgcgtc tctgttgctt | 1320 |
| | ctttgggtct cggagacctc aaccttttct tcagattgca aaccttcttg ccttcaagcc | 1380 |
| 20 | tcggctccaa caccagtccg gcagaggaac ccagtctaag gaggtacgct cccttctgc | 1440 |
| | cattctctat tccattaacc tgtttcgtgg taaacgtagg actgatcctc caaaattacc | 1500 |
| | ttattaatta gttacatat ttattatcta tctgtccac cagaatgcag gtttcggaa | 1560 |
| | ggcagggatt taaaaaaatc tgtttgttcc tatgtgattt tcccatacca agcaccgtgc | 1620 |
| | ccggcacaag ctgggatccc agtacacatc tcgggacgga agaaccgtgt ttccctagaa | 1680 |
| 25 | cccagtcaga gggcagctta gcaatgtgtc acagggtggg cgcccgcggt ccgggcggac | 1740 |
| | gcaactggctc ccggccgggc gtgggtgtgg ggcgagtggt tgtgtgcggg gtgtgcggg | 1800 |
| | tagagcggc cagcgagccc ggagcgcgga gctgggagga gcagcgagcg ccgcgagaa | 1860 |
| | cccgagcgcg cgccctggca gggcagctcg gaggtgggtg ggccgcgccg ccagcccgct | 1920 |
| | tgaggggtcc ccattggccg cctgccggcc gccctccgcc caaaaggcgg caaggagccg | 1980 |
| 30 | agaggctgct tcggagtgtg aggaggacag ccggaccgag ccaacgccgg ggactttgtt | 2040 |
| | ccctccgcg aggggactcg gcaactcgca gcggcagggt ctggggccgg cgcctgggag | 2100 |
| | ggatctgcgc cccccactca ctccctagct gtgttccgc cgcgccccgg ctagtctccg | 2160 |
| | gcgctggcg ctatggctcg cctccgacag cgctccggag ggaccggggg agtcccagg | 2220 |
| | cgcccggtg agtagccagg cgcggtccc cggtccccc gaccccggc gccagctttt | 2280 |
| 35 | gctttccag ccaggcgcg gtgggtttt tccgggcagt gcctcgagca actgggaagg | 2340 |
| | ccaaggcgga gggaaacttg gcttcgggga gaagtgcgat cgcagccggg aggcttcccc | 2400 |

agccccgcgg gccgggtgag aacaggtggc gccggcccga ccaggcgctt tgtgtcgggg 2460
cgcgaggatc tggagcgaac tgctgcgcct cggcgggccc c 2501

<210> 44
5 <211> 2501
<212> DNA
<213> Homo Sapiens

<400> 44
10
gatgtgaaaa gagaaataat tgaaaaagac tggagtacat atactatcta cagtgtctgt 60
tttaaagaaa caacattcta gcacacctt ctaccttga ctaagattac tgtaatgaga 120
gcaccagtac ccctgagtaa ccgaaagggc attttggaaa ctgagctttt ggtgtttata 180
tgaacattct gtcttcagg acctgccttg atttattcaa gactcatact gctgtatatg 240
15 gtgttgata cattaggggt agttgggtag cagtaactga tatagaaaat tttaaagtga 300
aaaaacactg gggagtgaac ctttccatta tatatatata tatatatata tatatatata 360
tatatatata tatatatata tataaattca catcaggatg agtttctgtt taggcaatgt 420
tggaanaacgc tatttccatt tttttttttt aacaaatatt taacaaacat ttataaggca 480
cttaaatcca tgctggctct tacaagtgtt gactcatttc tcataaccac cttggggtag 540
20 aaacggagag gctaaacaat ctgcaggcga tgcttacta ctaaatgcag gtggcagcct 600
tgccctgtgt ctctgcttg ctggaacac aggtcttacc tattgagctg ggctgtgtag 660
aactctgttg tggagacatc tggccctggg gcagaagcct ctgctttttc cccctcctcc 720
catcttactc catgtctcag agagctctga atcccacttg gagaatcaca cttaaacctc 780
ctaaaaacct aatgatgaat aaaaataagt tctctagaac ttctggagaa aaaagtaata 840
25 aagctaccag gttaaatgac tgaaattcct gagagaaaac aacatgtgtg tgtttctcta 900
gaaagggggc ccaatactga ataccaggaa gtcctatagt aaatggaatg tgactctatg 960
tgggatcccg cgttcctatt tcatccgaat gcatgtctgc tgcttcagtg ggaagggtgc 1020
ttgcacacca ggtaccact ccctgggtgc atgtgctatg cagtccaaag acagaaccag 1080
gaatggtgag cccatgagcc tgctggaccc agcccctccg aggtccggag tgacaaccag 1140
30 tgccgtatct ctgatcaaa cctgaacccc tcctacaggg aaaagatttc caggggattt 1200
tgaaagtcc aacattttac agggagaag gaagataagc aggatatgaa agaagagttc 1260
atgttataca gccctggctt ccactgacgc taacactgga ttcagctttt gactactgata 1320
atctgttgcc accaaatgga aaacgtaaac aagatattct aagtgtggtt agagaatatg 1380
caacacaagg aacaagcaga acattcttct ctggaatctg acataatgga ctgtactttc 1440
35 acagacagca ctgatgttag atgtacgtga aataggctaa actgaaaata agaaaggctg 1500
aggcagagag gataatatag ctccagccta tctccagca ccttggttaa tttctcaac 1560

ctccagccac aaatccgaga cacaacgctc ttcctccaaa gaggtcgcgc etttctctgtg 1620
 gtggtttctca gggatccgcc ccagctcctt ctccgttccc agccccacac actgggatca 1680
 ccaggcacc cagatccac ctctcaggtg gtatcttcag cgcaggctgc cactcagccc 1740
 ccctccaggg atctggggca gaaggcgaat atcccagagt ctccagagtc acaggagtta 1800
 5 ctctgaaggg cgaggcgagg gctgcatcag tggaccccca cccccaccc gcacccaag 1860
 cgctccaccc tggggcgagg gccgtcgctt tccttcggga ctccggatcg atctggaact 1920
 ccgggaattt ccctggcccg ggggctcccg gctttccagc cccaacccatg cataaaaggg 1980
 gtctcgaggat ctccgagagc cacagagccc gggccgcagg cacctcctcg ccagctcttc 2040
 cgctcctctc acagccgcca gaccgcctg ctgagcccca tggccgcgc tgctctctcc 2100
 10 gccgccccca gcaatcccg gctcctgcga gtggcgctgc tgctcctgct cctggtagcc 2160
 gctggccggc gcgcagcagg tgggtaccgg cgccctgggg tcccggggcc ggacgaggct 2220
 ggggtaggca ccagcgcgg acagcctcgc tcagtcagtg agtctcttct tccctaggag 2280
 cgtccgtggc cactgaactg cgctgccagt gcttcagac cctgcaggga attcacccca 2340
 agaacatcca aagtgtgaac gtgaagtccc ccggaccca ctgcgcccac accgaagtca 2400
 15 tgtaagtccc gcccgcgct gcctctgcca ccgccgggt ccagaccct cctgtgccc 2460
 caaccctgtc ccagcccgga cctcctgcct cagagattc c 2501

<210> 45

<211> 2501

20 <212> DNA

<213> Homo Sapiens

<400> 45

25 ggcgacagag caagactccc tcttaaaaaa aaaaaaaaaa aaagattctg agtcaaagtg 60
 ctcaagttga atgcattttg tcatccacaa gacaaatcgt gttaaccctt tgtggtttac 120
 tttatctata aaatagagat aacaatagtt cctgcttcta gggttgttgt gggaattaaa 180
 gacttagaat aatgttcagc ctctaatacag tgctgtcaca actgtctgat acaattgtat 240
 tatattttgtg tactttgtag attgatatta aatcatactt ttaaaaatag gtgcttaatg 300
 30 ttccactcaa ttaccttaaa acatgtttta ttatgtctct atcctactct tataacactt 360
 ctataaaaac tttttacata tagcgtccac ttttggttca gtttcttagg aaaataactt 420
 tgagagtcag ctatctgaac caaagaaaca ttaacattac cagactatat tgggattttt 480
 gagactggct tttatcaatt ctttagctac gggctcttgt catcatctct accagtgacc 540
 taagtgtcaa acccaaatgc cttgtatctg tcccattaaa gagatgcagc atctgctcct 600
 35 ttcttactgt ttccatttcc tctgcatgc ctctcttac aaccataaat atccaggtct 660
 cttaggtttt aaacggggca tctctcaacc cccacattct tttccttggt tattcccttc 720

cctccaacag ttcaattcac ctagatcccc acgcctgaaa ttatcctaga tgcctagag 780
gcgcctcatc attacaatgg tacattatc tccactcctt tacatgtcac gccagctttc 840
aaactgaaaa tctgagcgtt catccctggg gcatcacctt taaattccag atctccaaaa 900
tccagggtca tgtaacctta aaaaattttt accctctctt ctccactgcc ctgttcagg 960
5 ccttatctct tccagcagct gttccaaagg cctactctgt ttccctttcg gagtgcctaac 1020
ctccaccgaa gcctccaccc agttgccaat tctgccccat gcctgataat ttgctcgtgc 1080
gttgacatac ataaaatttc taagacaaaa attttttaat aatggtaaataaaccttg 1140
aactgcatac agatcataca gatccataat aagagaaaag gtcccagatt aacacgaaa 1200
actttccatt taactaacat ttgactgggt aaacttcac aagcaagacc ctacttaata 1260
10 ccacattacc ttctactgaa gaggttggtg tcattctctg gaaatatctg aattcattcc 1320
tacaagttag agaaacagcg ttactcgaaa cattatccct tgggctcgag ctctaaggca 1380
cctgacaaac ggagcgtgtt gggtaggggt gaggtgtttt ctccagggtt gggactttgc 1440
cctgggcgag ggcgccgag ggcaaagacc tcaccgggca gcagaatccg ggcagaaatc 1500
agcaactggg cctcccgcg agcagaaaag gggaatccag tcggggccca cccttctgc 1560
15 cagcgagac cgcaagtctg gccccatcct ctgcgggga gtcggcctgg cgcgtccgc 1620
ccaggtaacc cgaccgtggg cagcctgcgc ccgtttgggt cccatcgccc cggcccggca 1680
gatacctgag cggtgccag ggcaggctcc cgttcttgcc gatgcccatt ttctgggaca 1740
cagcgacgat gcagtttagc gaaccaacca tgacagcagc gggaggacct ccagcccg 1800
tcgttacagc agaacgcgc gtcaagttt gcgcgaaatt gtggccgccc cggccctcg 1860
20 tccctatttg tgcaggcgag gccccgccc ccgcccgg cgacgcagg gtcgaggcgt 1920
gtcgcgccc gcagacgctt gggaactgcg gccgagggtt cgcgtctctc gccgggacct 1980
gccgcccggc tgccatcctt gccctgccat gtctgcggg aagcctgcgt cggcgggcct 2040
cgctgcctcc agctcagccc ctgcgaggca agcgggtttg agccgattct tccagtctac 2100
gggaagcctg aaatccacct cctcctccac aggtgcagcc gaccagggtg accctggcgc 2160
25 tgcagcgccc gcagcgccc cagcgcccgc cttcccgc cagctgccgc cgcacgtagt 2220
aggttctgtc tgggactggg cagggccatc ggggctgggg gggcggggct tgtgggtaag 2280
gcgggcgag gcgtggacc tccgcccgat gatagggtg gaggaggaag gggcgggctg 2340
aagaagggga aggtgggaag agcccagccg gggtacaaa ttgggtgaag cgctgaggtt 2400
ttagtacttc cgtttgagga gataggcaaa ggttatgcag gtttttaatg gcaggcctga 2460
30 gacaggaact caggtctcct gactccatt ctgatgagg g 2501

<210> 46

<211> 1092

<212> DNA

35 <213> Homo Sapiens

<400> 46

| | | |
|----|---|------|
| | aagcttcccc ttcacatccc aagaaggcat tcaggtcttt ctgtgctagg ccccaggtaa | 60 |
| | agtgtctggac taccagtaa ttgggttcag tagcaggatg gcctcagatt gaggtcccag | 120 |
| 5 | ggccaaagga cactcctct cctcagcgt ggtccgggaa aggcaagctc cgggcgggag | 180 |
| | cgcacgcgc gcccccgaag cctggtctcc tcgccacgcc cacttcctgc cccatccc | 240 |
| | cgcctttcca ggtcttctcc cgtgaaccg gatgctctgt cagtctccta ctctgcgtcc | 300 |
| | tcggccgcgg cccgggtccc tcgcaaagcc gctgccatcc cggagggccc agccagcggg | 360 |
| | ctcccgagg ctggccgggc aggcgtggt gcggttagga gctgggcgcg cacggctacc | 420 |
| 10 | gcgcgtggag gagacactgc cctgccgcga tgggggccc gggcgtcct tcacgccgta | 480 |
| | ggcaagcggg gcggcggtg cgtacctgc ccaccgggag ctttcccttc cttctcctgc | 540 |
| | tgctgtgct ctgcatccag ctccggggag gacagaagaa aaaggaggta gaatggatcc | 600 |
| | ccttggcctt cccctgtggt cgggggcggg ccagggtggg ccgcgttgcc caggcagccc | 660 |
| | tgccgtgttg ctaggcagcc tggtcgccg cgtgggcgat gcggcgctg gggcgggagc | 720 |
| 15 | cgcgagggtg ggaggccctg gggcgtttcc gggacgtgga gttagcagg ttctgacttg | 780 |
| | aaaaacgacg gcaaagcgtg ttcttgactg cttctgagca cctcacacct ttcagaccca | 840 |
| | gggcgccttt attcccagct ggaagcccag cttagagcaa tgggtccact aaaaggggtg | 900 |
| | tgttgatgt gaaaataccc ttggaagta ttataagcc tcaggaat atgttttcct | 960 |
| | tattttctta ctctgtccc ttcattacc atttcaagaa gcaacagaac ctgtgcagag | 1020 |
| 20 | tgtgttttaa gttacactgt atgtttatt ttgtttatgt tgaactcgt gtatacttgt | 1080 |
| | gagaataagc tt | 1092 |

<210> 47

<211> 2501

25 <212> DNA

<213> Homo Sapiens

<400> 47

| | | |
|----|--|-----|
| 30 | cgaaatgaaa cctcgcccag gaggcgcgg acctggacac ccggcgccac ctcttcacc | 60 |
| | tctgaccag gtttctccc ggcgtgcga gctccgggg aagggttaga gccggcagcc | 120 |
| | ctcccagcc cggggagggg agagggttat gcgacccac ctctggctag ggccgggag | 180 |
| | gcctttgctt cccgggagcc ctgcccggc tccttggtcg cagggtgct gggteccag | 240 |
| | caggaacgag agggtaggc ccacatgtg cccggcgcc caggcggtg tgcagcgtcc | 300 |
| 35 | tcactgtccc ggctgccag ggctgcggc acgcggccag tcagcagcga gttcaggtcg | 360 |
| | cgcagatattt attgatgagc tctgactttc agcactttcc ctaagtcaag aagagtctag | 420 |

| | | |
|----|---|------|
| | cgtacccttc ggctgcttca tttcagcctc cctgcctcag ctcttcagcc ctattccccc | 480 |
| | tcgccctgtc ctgggggtgtg tacagcagcc caggccttcc ttctccttcc cggctccgtg | 540 |
| | gcccgaagcc gccgagagag ctcgggacag cgcaggacca ggcagccgct cgctctcctg | 600 |
| | tcaccttaac tgcaggctcc gaggggcgcc tttggagtgt actgaggtgt gtcctaatacg | 660 |
| 5 | tgcggcattc aacaaatgga cttctggtgt gtggtcagaa gagaaaagcc atttacttac | 720 |
| | tttctcctccc ggtttttctgg caacagctga aggggagctg cctccgtgga ctgagcagac | 780 |
| | ccaggagagg gagtcgtggt gcggagacac acgcaccaca cacagatgac cgggtggcaca | 840 |
| | cacgacacac gctgacatac cgacatcgcc agtgggacac acacacacac acacacacac | 900 |
| | acacacacac acacagagag agagagagaa tccctcccag catttggtcat ccgccccccc | 960 |
| 10 | accaggtct ccaactcccc tcccctctta tctcccctgg ctcccctcc tctcgggcgc | 1020 |
| | tgcgaaaagc agccgcactt agtcaacaaa tggcacgtgg gagaagttgg tgagtgtttg | 1080 |
| | gtgaggactc ttcagggtctt ttcacaagaa ccctctgtac acaaagtaag tggcgtgttt | 1140 |
| | actcgggcct ctccagccag agctgtgcct ctgctccgct gcgcaccgcg gcttccgaaa | 1200 |
| | ggagaaagga gagaagaaa ggcgggggaga gcgggggtgga ggatttgac aggccctgga | 1260 |
| 15 | ggcttgggct ggggaggcct ctggcctcgt ttagttctcg gcccggaac ctctctcgg | 1320 |
| | cctaggtctt gccgcggcct ccgcagctgg aatggagctg ccaggacca gtgacgtctc | 1380 |
| | cgcccccttc ctcttcttcc aaggggccag gtgggctggg gtgcggccgc cgctgtgctc | 1440 |
| | tgtgtcttgg ggccccggtt gggatgggtt gggggcgggc gggggcgggg cggcaggcca | 1500 |
| | cgctgtcctg gagttggcaa gaaaggacag cacagaaact tgcacctcc gaggactggg | 1560 |
| 20 | agtcccgagt ccagcttagg gggagtgggg gcgcgacccc caaccagaa accttcaatt | 1620 |
| | gaccgtcaa gtctcgcgca gcagggcggg ccgcgccgaa tctcggcgtg cgcggagcgg | 1680 |
| | ggagatgcag gcgagcgcca gagcccggtc tcgggggccc tgcgcggggg agaggagccg | 1740 |
| | ggaccaccg gcggagccga aaacaagtgt attcatattc aaacaacgg accaattgca | 1800 |
| | ccaggcgggg agaggagca tccaatcggc tggcgcgagg ccccggcgtt gctttgcata | 1860 |
| 25 | aagcaatatt ttgtgtgaga gcgagcgggt catttgcatg ttgcggagtg attagtgggt | 1920 |
| | ttgaaaaggg aaccgtggct cggcctcatt tcccgtctg gttcaggcgc aggaggaagt | 1980 |
| | gttttgctgg aggatgatga cagaggtcag gcttcgctaa tgggccagtg aggagcggtg | 2040 |
| | gaggcgaggc cgggcgcccg cacacacaca ttaacacact tgagccatca ccaatcagca | 2100 |
| | taggtgtgct ggctgcagcc acttccctca ccacactct ttatctctca ctctccagcc | 2160 |
| 30 | gctgacagcc cattttattg tcaatctctg tctccttccc aggaatctga gaattgctct | 2220 |
| | cacacaccaa cccagcaaca tccgtggaga aaactctcac cagcaactcc tttaaaacac | 2280 |
| | cgtcatttca aaccattgtg gtcttcaagc aacaacagca gcacaaaaaa ccccaaccaa | 2340 |
| | acaaaactct tgacagaagc tgtgacaacc agaaaggatg cctcataaag gtgagtcgcg | 2400 |
| | ttctttcttc tcgctttatt ttattgcaa tattcagaca ggtctcccc ttctctcccc | 2460 |
| 35 | cttcttctct cccctctcgc cgggtccctc cccactgct a | 2501 |

<210> 48

<211> 2501

<212> DNA

<213> Homo Sapiens

5

<400> 48

tgatggttgc acaactctga gtacatgaaa aatcaatgaa ctgatacttt gagtgagctg 60
tatgatactg gaattacacc tcaataaagc atggtaactg ttttaagata ggctggaaag 120
10 agaaagcctg aaaacaacaa taatgatatt aataaattag tttacttctc tagtctcata 180
tacttctgtg ccacacttg ctctgttct attcataatg gtccccctgc agttgccata 240
ttatatectg ccatttgatg ccgggtgaac attctatacc tgcttcccag aattctcttt 300
acctttctc tatctgccta acttccacat atctaaaatt aatcagagta aactatttac 360
tagaacaacc aactccaaat ctagtaacc taacatgata aaggtttggt tctcactcat 420
15 atagcccctc ccagatgat cgaggggtcc aggtcctta cctctagtgg ctccccacc 480
ttctggagtc ttctgcattc ttatacatg gttgagataa actatgagtc attagcacag 540
ctagaccttg aggtctaca agaaaattg caaatcattc actctgtttt gaacaaggta 600
tatttaagat gatgttaaaa taccatggg tcttgggtca aatacagttt atgactgtgt 660
atctaaaata tatattgcaa tattcttccc ttttctact gacttcatga atttagcggg 720
20 gatccatttt ataagctcaa agataattac ttttcagact aagaatattt agggtaaaaa 780
gtactgttca acatctctac tgaggatgtt atgatgtagc acactgtata agctggagct 840
aaaggaaact ttccttaaag tgctatttac taaaaattg aacacattcc ttaagacaaa 900
tcgaagtgtg gcacacaaca tccaaacttc catcatagat acagaggtgt taccatctcc 960
cactccaaa tttctttgtc acgtgagga tactcaagag gagcaggaca tgttggtcgc 1020
25 agcaggagaa acttgaaagc attcactttt atggaactca taaggagag aatttcttat 1080
tttagtatcg tccttgatac atttattatt ttaaaagata atgtagcaa atgtcttctt 1140
ctgtgttaaa tctttacaaa actgaaatct taaaatggtg acaaaaattc tacttctgat 1200
agaatctatt catttttcca attagatagg gcataattct taatttgcaa aacaaaacgt 1260
aatatgctta tgaggttcca tocaaagaa cctgctattg agagtagcat tcagaataac 1320
30 ggggtggaaat gccaaactca gagtttcaga tcctaccggt aattggggta gggaggggct 1380
ttgggcgggg cctccctaga ggaggaggcg ttgttagaaa gctgtctggc cagtccacag 1440
ctgtcactaa tcggggtaag ccttgttgta tttgtcgtg tgggtggcat tctcaatgag 1500
aactagcttc acttgtcatt tgagtgaat ctacaacccg aggcggctag tgctcccga 1560
ctactgggat ctgagatctt cggagatgac tgctgccgc agtacggagc cagcagaagt 1620
35 ccgacccttc ctgggaatgg gctgtaccga gaggtccgac tagccccagg gttttagtga 1680
gggggcagtg gaactcagcg agggactgag agcttcacag catgcacgag tttgatgcca 1740

| | | |
|----|--|------|
| | gagaaaaagt cgggagataa aggagccgcg tgctactaaa ttgccgtcgc agccgcagcc | 1800 |
| | actcaagtgc cggacttggt agtactctgc gtctccagtc ctccggacaga agttggagaa | 1860 |
| | ctctcttggga gaactccccg agttaggaga cgagatctcc taacaattac tactttttct | 1920 |
| | tgccgtcccc acttgccgct cgctgggaca aacgacagcc acagttcccc tgacgacagg | 1980 |
| 5 | atggaggcca agggcaggag ctgaccagcg ccgccctccc ccgccccga cccaggaggt | 2040 |
| | ggagatccct ccggtccagc cacattcaac acccactttc tcctccctct gccctatat | 2100 |
| | tcccgaaaacc cctcctcct tcccttttcc ctctccctg gagacggggg aggagaaaag | 2160 |
| | gggagtccag tcgtcatgac tgagctgaag gcaaagggtc cccgggctcc ccacgtggcg | 2220 |
| | ggcggccccc cctccccga ggtcggatcc cactgctgt gtcgccagc cgcaggctcg | 2280 |
| 10 | ttcccgggga gccagacctc ggacaccttg cctgaagttt cggccatacc tatctccctg | 2340 |
| | gacgggctac tcttccctcg gccctgccag ggacaggacc cctccgacga aaagacgcag | 2400 |
| | gaccagcagt cgctgtcgga cgtggagggc gcatattcca gagctgaagc tacaaggggt | 2460 |
| | gctggaggca gcagttctag tccccagaa aaggacagcg g | 2501 |
| 15 | <210> 49 | |
| | <211> 2501 | |
| | <212> DNA | |
| | <213> Homo Sapiens | |
| 20 | <400> 49 | |
| | taccttcata aaaggatctt tgacttggtg agtgtgtgct atgcatactt ttcagtgtac | 60 |
| | accacaagtgc ccacttagca actccactag acagggcagt gtttcagcat ggggtggggt | 120 |
| | gccccctgac aggcttttaa aaggcccgga tgccaatgca cattccaaca ctatccacaa | 180 |
| 25 | aaaggagact ggagcagtgc tcttccctgc attgggcaag gagactctcc ctccctgcct | 240 |
| | aaccacttgc ctgccctggt ttgtgggaga attacaagta aatgctacag aggcagtgga | 300 |
| | gaaaaaaggg tgttttaatt cctctccaga gtttccttta tttgatgtat gttgcatcct | 360 |
| | ttaaacaagt tgtgcaaaat ggctgcaggg tagattggct ctccctttta aagctctcca | 420 |
| | tccggctggg tttatttgta aatactgcat ctatccttct tagtgtttta ggactggctg | 480 |
| 30 | gaaagactct tcttctgta ggttgggtca gtgtgagaga tctaaaaaat cattttccct | 540 |
| | taaaattact gtattttaat aaaaggattg ggcaggggct ggaatgagag aaaactggtc | 600 |
| | cttcaaaatg taaaactgtc atacttaaac cagtttacia aatatgcgtt taattatgtg | 660 |
| | gtgggatgtg tgtagtgta tgatgagaga ggcaaccaac atggctatgt ggggtgcaag | 720 |
| | gatgtgggaa caggcaagta attttcacat tggactttca tcctaggagag ctgggttcta | 780 |
| 35 | gtcacagctc tgagctgtgt gaccttgggt aggtctcatc tccccgggtt tttgtttcac | 840 |
| | cagttgaaca gtatgaggat gagtacagc taacatttgt tccatgatat ttaccagca | 900 |

ccatacaagt gttatttctg tcctcccagt taacactgac gtgggtagta ttatatgccc 960
 atttttacaga tgaggaaact gaagcctgaa gaagttaaact acttatccca gaacacacag 1020
 ctggtaagtg gcagacctgg aattggaatc tagttcagtt tgattcccca acccatgctc 1080
 ttgaccacta tactgttttt tcaagtccag atctgaaatc tcattttctg tgtggctgtg 1140
 5 tgtttgggac aggggtaacc aattcctgac tactctatat gctgcataga acctggagag 1200
 gatTTTTcaa agtaaatgaa tctcgaaagc tggattgcag agcaaacgag tgcagtcaat 1260
 tcagccaggg gcttgcaaga gggagaaaga gaaaagact gtggaatgga aagtttccca 1320
 acccaagcct ttccaaggg gtagccattc tctgttctac agtttagggc ttgcatgtgc 1380
 tttttctgga gtggaaaaat acataagtta taaggaattt aacagacaga aaggcgcaca 1440
 10 gaggaattta aagtgtgggc tggggggcga ggcggtgggc gggaggcgag cgggcgcagg 1500
 cggaacaccg tttccaagc taagccgccg caaataaaaaa ggcgtaaagg gagagaagtt 1560
 ggtgctcaac gtgagccagg agcagcgtcc cggctcctcc cctgctcatt ttaaaagcac 1620
 ttcttgtatt gtttttaagg tgagaaatag gaaagaaaac gccggcttgt gcgctcgctg 1680
 cctgcctctc tggctgtctg cttttgcagg gctgctggga gtttttaagc tctgtgagaa 1740
 15 tcctgggagt tgggtgatgtc agactagtgt ggtcatttga aggttagcag cccgggtagg 1800
 gttcaccgaa agttcactcg catatattag gcaattcaat ctttcattct gtgtgacaga 1860
 agtagtagga agtgagctgt tcagaggcag gagggctctat tctttgcca aggggggacc 1920
 agaattcccc catgcgagct gtttgaggac tgggatgccg agaacgegag cgatccgagc 1980
 agggtttgtc tgggcaccgt cggggtagga tccggaacgc attcggaagg ctttttgcaa 2040
 20 gcatttactt ggaaggagaa cttgggatct ttctgggaac cccccgccc ggctggattg 2100
 gccgagcaag cctggaaaat ggtaaatgat catttgatc aattacaggc ttttagctgg 2160
 cttgtctgtc ataattcatg attcggggct gggaaaaaga ccaacagcct acgtgccaaa 2220
 aaagggcgag agtttgatgg agttgggtgg acttttctat gccatttgcc tccacaccta 2280
 gaggataagc acttttgagc acattcagtg caaggagat catgtttgac tgtatggatg 2340
 25 ttctgtcagt gagtctggg caaatcctgg atttctacac tgcgagtcg tcttctgca 2400
 tgctccagga gaaagctctc aaagcatgct tcagtggatt gacccaaacc gaatggcagc 2460
 atcggcacac tgctcaatgt aggtttattt ttttccttc t 2501

<210> 50

30 <211> 2501

<212> DNA

<213> Homo Sapiens

<400> 50

35

ggaggataga aatataaatt aaagaatgac acaaataatt ataaagttac agctgttaaa 60

| | | |
|----|--|------|
| | agaaaaagcat atggtgccaa gagaacgtgt aatacaagat ctactcatgg aggtgagggg | 120 |
| | aagcttgccc atcaaagaag ttatgattca atccacgaag accaggagtt ggctgggtga | 180 |
| | agaaaaaaag gtcagaggaa ggaagtccac actggggaag gctctaagca taaagggtag | 240 |
| | gaggattaca gaggcatatt cacgaaatth ggagaaggct ttcagtaagc aaggagaagc | 300 |
| 5 | caaatgaaag ttacgggag agttggaggc ttgaagacac gttcaaggat ctggttttta | 360 |
| | tcttctcttt atctcaagag cagtgggaag ccattaaatg attttaatca gagggttggt | 420 |
| | ataactagtt ttgtatthtg aaaagctgaa ttcagctctc gtttgagaaa ctgagtgaag | 480 |
| | gagcccagaa cggccgtggc tgagggtgac tcgtgggaga ctctacaca agccatggca | 540 |
| | gtggcatggg ctggtggcag aagagggaat agggagaaga tttggaactc aatcttctc | 600 |
| 10 | cattgacaaa gtcactccag ctttggcaag gcaattaatt ggtgggaaag aagatgccta | 660 |
| | gccctcctga tttcactgca ctttctgcat ctcaacatg agtactggga agtggcaaaa | 720 |
| | catccagagg cagcttgggt gctaggtgga gcatgagtha aaattccagg atgaagcaaa | 780 |
| | tgaacactta gaatgacagg aaagatttgg gagtgggtt tgggggaggg ctatttacct | 840 |
| | ttattccctg gagaccctgg cacaaccct tgcctctgca atcttctct caggtaaagg | 900 |
| 15 | aattcattaa atgaattgct agaagatcta ctgaccagag ggctgtacag aatcatatct | 960 |
| | ttgagagtgg gaagtaggtt gatcacatag tttattatcc aatcaggaca tatctgaaag | 1020 |
| | agaaagggg ttctattaat atttaacta caaacatgt acaccaggaa tgtcttggg | 1080 |
| | aaatctggtt gccctagcaa gaaaggaaat ttgaaagttt atactgttct gctcccatgt | 1140 |
| | taccccgtht gcacatgaga gggtaagtat tctcttctt cacctgcatt aagggaataa | 1200 |
| 20 | aagcacaagc attcaggtga ctcccaaccc actthtaatt ttacagthtc tgctatactc | 1260 |
| | tatacattct gaaaattaca tttccacca ctatcacttc gtgatagggtg atcatttaca | 1320 |
| | attactcact gactcagthc cggaagagg cggthcaaaa tgggacgctc tatccaggth | 1380 |
| | ctcattagaa atgcagaatc tctgcctgcc tcttagacct actgaattag aatctgcatt | 1440 |
| | thtaataag atttccaggt gatcaatatg tacattaaaa cttgagaaaa acctctagac | 1500 |
| 25 | ttcgacctaa agaaaaacat thtaactt gacagtgtat gcacatacat acatgcatat | 1560 |
| | agacacaact gaagcacaaa thtaatgaag tagaatttac cgttactatt ttatttggga | 1620 |
| | aagaaatgtg ctgcgactc aatagattgg agtattcact cctggatctc aacttgcaat | 1680 |
| | ttgaaaacgc atctctaaag cacctaggag caatctgaag aaagctgagg ggaggcggca | 1740 |
| | gatgttctga tctactaggg aaaacgtgga cgtthtctgt tgttactthg tgaactgtgt | 1800 |
| 30 | gcacttagtc attcttgagt aaatactthg agcgaggaa ccttgagtgg tgtggagggg | 1860 |
| | cggtagggg cagctgaaag tgggcaaa cctcggagg ggctggtcta ggaaacatga | 1920 |
| | ttggcagcta cgagagagct aggggctgga cgtcaggagg agggagaagg ctctcggg | 1980 |
| | gagagaggth ctgcccagct gttggcgagg agthtctgt thccccgca gcgctgagth | 2040 |
| | gaagttgagt gactcactcg cgcgcacgga ggcacgacac ccccgcgct gcacccgctc | 2100 |
| 35 | gggacaggag ccggactcct gtgcagcttc cctcggccgc cgggggcctc cccgcgctc | 2160 |
| | gccggcctcc agggccctc ctggctggcg agcgggcgcc acatctggcc cgcacatctg | 2220 |

| | | |
|----|--|------|
| | cgctgccggc cggcgcggg gtccggagag ggcgcggcgc ggaggcgag ccaggggtcc | 2280 |
| | gggaaggcgc cgtccgctgc gctgggggct cggctctatga cgagcagcgg ggtctgccat | 2340 |
| | gggtcggggg ctgctcaggg gcctgtggcc gctgcacatc gtcctgtgga cgcgtatcgc | 2400 |
| | cagcacgata ccaccgcacg ttcagaagtc gggtagtggt tccccagccc gggctcggcg | 2460 |
| 5 | gggcgcgggg ggtcttcctg gggccccgc ctctccgctg c | 2501 |
| | <210> 51 | |
| | <211> 2500 | |
| | <212> DNA | |
| 10 | <213> Homo Sapiens | |
| | <400> 51 | |
| | ttcccatcaa gccctagggc tcctcgtggc tgctgggagt tgtagtctga acgcttctat | 60 |
| 15 | cttggcgaga agcgccctacg cccccctac cgagtcgccg ggtaattctt aaagcacctg | 120 |
| | caccgcccc cgcgcgctg cagagggcgc agcaggctct gcacctcttc tgcattctcat | 180 |
| | tctccaggct tcagacctgt ctccctcatt caaaaaatat ttattatcga gctcttactt | 240 |
| | gctaccacgc actgatatag gcactcagga atacaacaat gaataagata gtagaaaaat | 300 |
| | tctatatcct cataaggctt acgtttccat gtactgaaag caatgaacaa ataaatctta | 360 |
| 20 | tcagagtgat aagggttggt aaggagatta aataagatgg tgtgatataa agtatctggg | 420 |
| | agaaaacggt aggggtgtgat attacggaaa gccttcctaa aaaatgacat tttaactgat | 480 |
| | gagaagaaa gatccagctg agagcaaacg caaaagcttt ctcccttcca cccttcatat | 540 |
| | ttgacacaat gcaggattcc tccaaaatga ttccaccaaa ttctgccctc acagctctgg | 600 |
| | cttgacgaat ttccaccccc aaaatgttag tatctacggc accaggctcg cgagaatcct | 660 |
| 25 | gactctgcac cctcctcccc aactccattt cctttgcttc ctccggcagg cggattactt | 720 |
| | gcccttactt gtcattggga ctgtccagct ttgtgccagg agcctcgag gggttgatgg | 780 |
| | yattgggggt ttcccctccc atgtgetcaa gactggcgct aaaagttttg agcttctcaa | 840 |
| | aagtctagag ccaccgtcca gggagcaggt agctgctggg ctccggggac actttgcgtt | 900 |
| | cgggctggga gcgtgcttcc cagcaggtg acacgcttcc ctggattggg taagctcctg | 960 |
| 30 | actgaacttg atgagtcctc tctgagtcac gggctctcgg ctccgtgtat ttccagctcg | 1020 |
| | ggaaaatcgc tggggctggg ggtggggcag tggggactta gcgagtttg gggtagtggt | 1080 |
| | gatggaagct tggctagagg gatcatcata ggagttgcat tgttgggaga cctgggtgta | 1140 |
| | gatgatgggg atgttaggac catccgaact caaagttgaa cgcctaggca gaggagtgga | 1200 |
| | gctttgggga accttgagcc ggcctaaagc gtacttcttt gcacatccac ccggtgctgg | 1260 |
| 35 | gcgtagggaa tccctgaaat aaaagatgca caaagcattg aggtctgaga cttttggatc | 1320 |
| | tcgaaacatt gagaactcat agctgtatat tttagagccc atggcatcct agtgaaaact | 1380 |

5 ggggctccat tccgaaatga tcatttgggg gtgatccggg gagcccaagc tgctaaggtc 1440
 ccacaacttc cggacctttg tccttcctgg agcgatcttt ccaggcagcc cccggctccg 1500
 ctagatggag aaaatccaat tgaaggctgt cagtcgtgga agtgagaagt gctaaaccag 1560
 gggtttgccc gccaggccga ggaggaccgt cgcaatctga gaggccggc agccctgtta 1620
 ttgtttggct ccacatttac atttctgcct cttgcagcag catttcgggt ttctttttgc 1680
 cggagcagct cactattcac ccgatgagag gggaggagag agagagaaaa tgccttttag 1740
 gccggttccct cttacttggc agagggaggc tgctattctc cgcctgcatt tctttttctg 1800
 gattacttag ttatggcctt tgcaaaggca ggggtatttg ttttgatgca aacctcaatc 1860
 cctccccttc tttgaatggt gtgccccacc cccgggtcg cctgcaacct aggcggacgc 1920
 10 taccatggcg tagacagga gggaaagaag tgtgcagaag gcaagcccg aggcactttc 1980
 aagaatgagc atatctcatc ttcccggaga aaaaaaaaa agaattgtac gtctgagaat 2040
 gaaattttga aagagtgcaa tgatgggtcg tttgataatt tgcgggaaa aacaatctac 2100
 ctgttatcta gctttgggt aggccattcc agttccagac gcaggctgaa cgtcgtgaag 2160
 cggaaggggc gggcccgag gcgtccgtgt ggtcctccgt gcagccctcg gcccgagccg 2220
 15 gttcttcctg gtaggaggcg gaactcgaat tcatttctcc cgctgcccc tctcttagct 2280
 cgcggttggt tcattccgca gtttcttccc atgcacctgc cgcgtaccgg ccactttgtg 2340
 ccgtacttac gtcacttttt tcctaaatcg aggtggcatt tacacacagc gccagtgcac 2400
 acagcaagtg cacaggaaga tgagttttgg cccctaaccg ctccgtgatg cctaccaagt 2460
 cacagaccct tttcatcgtc ccagaaacgt ttcacacgt 2500

20

<210> 52

<211> 286

<212> DNA

<213> Homo Sapiens

25

<400> 52

30

tttgcactag gctggaagtg gccgccagtc ccccgtagaa ttccattctc tggaaaagtg 60
 gaatcagctg gcattgccca gcgtgatttg tgaggctgag ccccaacagt ccaaagaagc 120
 aaatgggatg ccacctccgc ggggctcgct cctcgcgagg tgctcacccc gtatctgcca 180
 tgcaaaaacga gggagcggtta ggaaggaatc cgtcttgtaa agccattggt cctggtcatc 240
 agcctctacc caatgctttc gtgatgctgc tgctgatcta tttggg 286

35

<210> 53

<211> 1400

<212> DNA

<213> Homo Sapiens

<220>

<221> unsure

5 <222> (1371)

<223> unknown base

<400> 53

10 ttccagctgt caaaatctcc cttccatcta attaatcct catccaacta tgttccaaaa 60
 cgagaataga aaattagccc caataagccc aggcaactga aaagtaaag ctatgttgta 120
 ctttgatcca tggtcacaac tcataatctt ggaaaagtgg acagaaaaga caaaagagtg 180
 aactttaaaa ctgaattta ttttaccagt atctcctatg aagggctagt aaccaaata 240
 atccacgcat caggagaga aatgccttaa ggcatacgtt ttggacattt agcgtccctg 300
15 caaattcttg ccatcgccgc ttcctttgtc catcagaagg caggaaactt tatattggtg 360
 acccgtggag ctcacattaa ctatttacag ggtaactgct taggaccagt attatgagga 420
 gaatttacct ttccgcctc tctttccaag aaacaaggag ggggtgaagg tacggagaac 480
 agtatttctt ctgttgaaag caacttagct acaaagataa attacagcta tgtacactga 540
 aggtagctat ttcattccac aaaataagag ttttttaaaa agctatgtat gtatgtgctg 600
20 catatagagc agatatacag cctattaagc gtcgtcacta aacataaaa catgtcagcc 660
 tttcttaacc ttactcgccc cagtctgtcc cgacgtgact tctcagccc tctaaagacg 720
 tacagaccag acacggcggc ggcgccggga gaggggattc cctgcgcccc cggacctcag 780
 ggccgctcag attcctggag aggaagccaa gtgtccttct gccctcccc ggtatcccat 840
 ccaaggcgat cagtccagaa ctggctctcg gaagcgctcg ggcaaagact gcgaagaaga 900
25 aaagacatct ggcggaaacc tgtgcgcctg gggcggtgga actcggggag gagagggagg 960
 gatcagacag gagagtgggg actaccccct ctgctcccaa attggggcag cttcctgggt 1020
 ttccgatttt ctcatctccg tgggtaaaaa accctgcccc caccgggctt acgcaatttt 1080
 ttaagggga gaggagggaa aaatttggtg ggggtacgaa aaggcggaaa gaaacagtca 1140
 tttcgtcaca tgggcttggt tttcagttt ataaaaagga aggttctctc ggtagcgac 1200
30 caattgtcat acgacttgca gtgagcgtca ggagcacgtc caggaaactcc tcagcagcgc 1260
 ctccttcagc tccacagcca gacgcctca gacagcaaag cctacccccg cgccgcgccc 1320
 tgcccgccgc tgcgatgctc gcccgcgccc tgctgctgtg cgcggtcctg ncgctcagcc 1380
 atacaggtga gtacctggcg 1400

35 <210> 54

<211> 2501

<212> DNA

<213> Homo Sapiens

<400> 54

5

10

15

20

25

30

35

gataatcttt tcatacaaga tgcattctgc ttttgtgggc ctcttgacgc cctcaagccc 60
ccatctgatt tgtacacaat gatccagtgg gccagaggag cccagagcca tgagcgggcc 120
atccctccaa gaactatttc tgactgtcca gtatcatgga gcaagtggaa agaagaaaaa 180
aaaaacccaa ttacttttcg aagagcaaga tgaatgctgt agaaggagaa ggaaggggag 240
ggagatggat ggggtccgat tccagaatct tcagatctgc ttggatgaat cattacctat 300
gatttgcggg acaagaatct gattttattc atcaaccagt agaaactttt ctttctgcct 360
cccaacatct gaaatccaac aaacatgtgc cttaggaaca taccggtcac cttttagagg 420
cattttatat acatattgag taactagaaa acactctttc cgtaatacac acacacacac 480
acacacacac acaccatctt gtcatacaac actcccacgc aagaaaagcg aaactgctgt 540
ttgatgaatg taaacacttg gctgtttgca gcagtcggga gtcctgccag gtttaagtgc 600
taagatggga ggtgaacccc aggggtttcc ccctgcccg tctgagatcc ttatttggtc 660
aagcttctac ctatgccctg gcctcggagc gagcccata gcgctggatc acagcagagg 720
gagcgaggcg gctgacgtcc catcccgaag agatgaatgg aattccagga agctagagtc 780
atgctggctt gggacagtgg cttggagacc agacttcaat gacagaagca ctaggcagcg 840
gcactcatgg caatgtgtgc acccacagaa atgtaacca cacctcgggt tcaggagccg 900
aaaaatgaaa agaacgttta gggaggaaaa agggaaatac aataataggc agagagtaat 960
ttattactct atgggtctgc tctgtaaata gctgaagact ctggagccag atggttctgc 1020
aaattctcca aacaggagtc acgttaagaa gcacgagtgg gcacaaaaac tgtttttcaa 1080
gacacaattt caatttggct tgtggaaact ggatacgagt aagtttcctt aaaattcgag 1140
tagaaagcag ctgtcctccc cgggccctt gatgagaata cgacacaccg cccaagcgg 1200
ccggccgagg gagcgccgag gcagcgggag aggcgtctct gtgggcccc tggcagccgc 1260
ggcaggaaaag ggcccgaagg cagcgaaggc gaacgcggcg caccaacctg ccggccccgc 1320
cgacgcgcg ctcacctccc tccggggcgg gcgtggggcc agctcaggac aggcgctcgg 1380
gggacgcgtg tcctcacccc acggggacgg tggaggagag tcagcgaggg cccgaggggc 1440
aggtacttta acgaatggct ctcttggtgt ccctgcgcc ccgtcggccc atttttcttt 1500
ttacaaaacg ggcccagtct ctagtatcca cctctcgcca tcaaccagge attccgggag 1560
atcagctcgc ccgaaagccc ctgcgccacc ccgcgggccc tcctaggtgg tctccccagc 1620
cccgctccctt ttcgggatgc ttgctgatca ccccgagccc gcgtggcgca agagtacgag 1680
cgccgagccc gtgcgcgcca aggctgcgtg ggcgggacc gacttttctg agaagttcta 1740
gtgctcccaa gccccagccc ccgccccctt cactttctag ctggaaagtt gcgcgccagg 1800
cagcgggggg cgagagagg agcccagact ggccccacc tcccgttcc tgcccggccg 1860

| | | |
|----|---|------|
| | ccgcccattg gccggaggaa tccccaggaa tgcgagcgcc cctttaaaag cgcgcggctc | 1920 |
| | ctccgccttg ccagccgctg cgcgcgagct ggcctgcgag ttcagggctc ctgtcgctct | 1980 |
| | ccaggagcaa cctctactcc ggacgcacag gcattccccg cgcctccca gccctcgccg | 2040 |
| | ccctcgccac cgtccccggc gcgcgcgctc cggtagacac aggtagtcg cccccggcg | 2100 |
| 5 | ccgcccaggga ccaaagctgc ccgggacatc cacctggagc gctgaggctt cagtcctct | 2160 |
| | ggtggacccc ggaacctaca ctctccccgc tcgcctaccc cagcccgctc ctctcagccg | 2220 |
| | ctggaggact cttcagggca aggtccaga gccatcctct ccagccttga ggttcacaaa | 2280 |
| | ccaactcatc aggacacccc aagatttcct tactctctga agtcctcctt aagcctttgt | 2340 |
| | atcagcactc cagggaagag tctgtacttc cctgcccctc cctgcaaccc caaactacag | 2400 |
| 10 | ttcctgatct tgctcacctt cgacttccca aaagccccca aattgttggc ctgcccccc | 2460 |
| | ccacacttta aaaccagcat ctctttctc cactctctc t | 2501 |
| | <210> 55 | |
| | <211> 7258 | |
| 15 | <212> DNA | |
| | <213> Homo Sapiens | |
| | <400> 55 | |
| 20 | ttcaatagga agcaccaaca gtttatgcc taggactttg tccccacaat cctgtaacat | 60 |
| | catatcacga cacctaacc aatccttatc aagccctgtc aaaaacggac tttaaaccaa | 120 |
| | gctgcaaatt ttcagtaatc tggccttgcc tttcccctc tgatagcacc atcaaaaaa | 180 |
| | cccccttact gccgaaagca ataagcccg ctttgttcca tccactgggt gtgttggga | 240 |
| | tatctgggga ctgccactga acagacgcac agaggagcc cctacaggca ggggttttc | 300 |
| 25 | tgtctgtgct tcttgggaga gtatgtctc tacatttgc gcgtgatgaa gacttcacag | 360 |
| | ctccatccag cgaccagact cacagctcca tccagctgc gcaaggggt ctgaggcagt | 420 |
| | cttaggcaag ttggggccca gcgggagaag ttgcagaaga actgattaga ggaccagga | 480 |
| | ggcttcagag ctgggctgag gtagagagtc tcctgtgcgc cttctctcct ctctgcaatt | 540 |
| | cggggactcc ttgcaactgg gcagccccg gcaggtgcat gggaggagc acggagaatt | 600 |
| 30 | tacaagcctc tcgattcctc agtccagacg ctgttgggtc cctccgctg gagatcgcg | 660 |
| | ttccccaaa tctttgtgag cggttcggaa gcacgcggg tccgggtcgc tgagcgctgc | 720 |
| | aagacagggg agggagccgg gcgggagagg gaggggcgg gcggggcg gacctgat | 780 |
| | agagcagggc ccgcgggtcg cagcacagtc ggagaccgca gcccgagcc cgggccagg | 840 |
| | tccacctgtc ccgcagcg gcgctcgc cctcctgcc cagccaccg tgagtgcgc | 900 |
| 35 | ggtcctgaga tccccgggc ggatgcgcg cgccccagc tccgagcgt ctgctgccc | 960 |
| | cgccctgggc tgcccggt cctgggtc cccggcggt gcacggagtc aaggcgcccc | 1020 |

| | | |
|----|--|------|
| | gtccccggcg tccccgcggt gtgccgatcc aggetgcccc gagtccggag cccatagagg | 1080 |
| | agagagacag ctggggagcc tggtcaccgc gggcatctcc cctgcgctgc agtcgccgc | 1140 |
| | ctggcctgcc tccccgttcc tccgcctctt gccctgactt ctccctctct tgagagccg | 1200 |
| | ccgtctagcg ccccgacctc gccaccatga gagccctgct ggcgcgctg cttctctgcg | 1260 |
| 5 | tcctggctgt gagcgactcc aaagtgagtg cgctcttctt ttgactgatg ctgcccgaagg | 1320 |
| | acctctgac agcaccaggg gagaggagg gctgctcagg gagctggggg ctccggattc | 1380 |
| | catccacagc agggccagac tctccccagg aaatgggaca ggggtggcagc ggaggcttga | 1440 |
| | gaaccacggg ggttggcact ggctggcaag ggaggaagag ggccaccggg actgccccag | 1500 |
| | cctgcgggca tctggtagat gaagcttaat ccatttctcc tggctggaaa ccatggtctt | 1560 |
| 10 | ccatttgaga actagatacg aacagggtga ggcgagaggg agagggaaga gtgggttttg | 1620 |
| | ggattggggc cagtttacc tccacctgga tccctggagc atgggacctt tgatgaagcc | 1680 |
| | tcctcccgaa tctcttccag ggcagcaatg aacttcatca agttccatgt gagtatccac | 1740 |
| | ccctacaaca gttggtgca cagacaagtt gggaaggctt caggggacac tccccctcct | 1800 |
| | gccctctgct gcagcgtgcg ccccccctta ccacttccac tccccctcgc ttaccccacc | 1860 |
| 15 | tttgttctct ccagcgaact gtgactgtct aaatggagga acatgtgtgt ccaacaagta | 1920 |
| | cttctccaac atttactggt gcaactgccc aaagaaattc ggagggcagc actgtgaaat | 1980 |
| | aggtatgggg atctccactg caactgggag agaaatttgg ggacaggag ggatgggtgg | 2040 |
| | gaggcaagag caggcaggag ttaggagctg gaggtagggt gggtagacac ttcattcccta | 2100 |
| | tgtgacaagc ataaacacac acacacgctc acgaaacagt ggccacacaa atgtgaggtg | 2160 |
| 20 | gggttggaag gagacctgt ccagtctctt ggcaggtctg aaacgacac tttaaaatgt | 2220 |
| | ccgttggcag cggggcatgg ttggtcacgc ttgtaatccc agcattttga gaggtcaagt | 2280 |
| | ttgagtggat catttaggtc aggagttcaa gaccagcctg gacaacatgg tgtaaccctg | 2340 |
| | cctctactaa aaatgcaaaa atcagcctgg catggtggtg gatgcctgta gtcccagcta | 2400 |
| | cttgggaggc tgaggcagga gaattgcttg aacatgggag gccagatctc agtgagctga | 2460 |
| 25 | gatcacacca ctgcactcca actgggcgac agagcaagac tccatctcaa aaaaaaaaa | 2520 |
| | aaataaaagt tagttggaat gttcttctct ttctcatatt ctctcatcct cctgtcccct | 2580 |
| | tgtagataag tcaaaaacct gctatgaggg gaatggtcac ttttaccgag gaaaggccag | 2640 |
| | cactgacacc atgggccggc cctgcctgcc ctggaactct gccactgtcc ttcagcaaac | 2700 |
| | gtaccatgcc cacagatctg atgctcttca gctgggcctg gggaaacata attactgcag | 2760 |
| 30 | gtgaggtggg ggcaacaagg accaaaagcc ctccctacag cttcccagaa accttggtac | 2820 |
| | catccccttc tcccagaggg ctggccatag cacaagagaa gtgcggcctc tggttgagtc | 2880 |
| | ttccctgagg ggaggaggca ggggaaggcc tctgggttgg aatgacatcc cctatcttct | 2940 |
| | tgtgttgtgc caggaacca gacaaccgga ggcgaccctg gtgctatgtg caggtgggcc | 3000 |
| | taaagccgct tgtccaagag tgcattggtc atgactgcgc agatggtgag catcactgac | 3060 |
| 35 | ctgctgatga cagggtgggtg gaaggggaca aacttcatg tccccttatt ccatcacag | 3120 |
| | aggactgagg aggtgggggg tgcccagag ggatgcttct tcctacctgc ctccctaaga | 3180 |

| | | |
|----|--|------|
| | catccctctg tttgtcctcc aggaaaaaag ccctcctctc ctccagaaga attaaaaattt | 3240 |
| | cagtgtggcc aaaagactct gaggccccgc ttaagatta ttgggggaga attcaccacc | 3300 |
| | atcgagaacc agccctgggt tgcggccatc tacaggaggc accggggggg ctctgtcacc | 3360 |
| | tacgtgtgtg gaggcagcct catgagccct tgctgggtga tcagcgccac acactgcttc | 3420 |
| 5 | atgtacggcc ctgggtttct cctcttcgac tctctgccc caccccaagc acatcccttt | 3480 |
| | ctccttccca gcaaagtgtt ccgcctcatt tctccctcat ctgcccctgt ccatgcgccc | 3540 |
| | atggccttgg ggacaagtcg tgctttgagg cctctaggga gggaaggaag aagtggcatg | 3600 |
| | atttcatggg actaagctgt ttgatgggta tcttcttcca cagtgattac ccaaagaagg | 3660 |
| | aggactacat cgtctacctg ggtcgctcaa ggcttaactc caacacgcaa ggggagatga | 3720 |
| 10 | agtttgaggt ggaaaacctc atcctacaca aggactacag cgctgacacg cttgctcacc | 3780 |
| | acaacgacat tggtagggg gaacgcccgc gactactgtg gccataatgg cttggggaga | 3840 |
| | gtgggaccca gggagagact ggagctgagt tgaagctgcc ggtggggcag ggtggggcg | 3900 |
| | agggaccttg aagcctcgat atacatgaca aaggatggca gggaagagtt ccatgaagtc | 3960 |
| | tgaggggcct ggtgctctc tggagagacc ctgaatttcc ccaacaagta gccctcttgc | 4020 |
| 15 | gagtggaaac agccctgttg gtatatggct tgggctggga aggcctgtt tatatgaatt | 4080 |
| | agaaaaagac acaccttctt ttgtgggatg cagcctctgt ctgtgctagg atatagaact | 4140 |
| | tggagaatgg agccttggga tggattccag cctaactacc tcagggggat cctctagagt | 4200 |
| | gcagctggga gtttttgag aaacgacctg tacagctgta tgcagtggct ctggccatcc | 4260 |
| | aagccttttt caacacctgg aacaaagccc ttggggcatg gggcagggga ggtttccagg | 4320 |
| 20 | tgataagcga ccagcagacc tccctggatg actgacctag ggataggcat agctacttcc | 4380 |
| | tccgcacttg gaggggacag atggggaccg cctaaccagt agtgatcttt ctctctgac | 4440 |
| | cctctgtcct cccccagcct tgctgaagat ccgttccaag gagggcaggt gtgcgcagcc | 4500 |
| | atcccgact atacagacca tctgcctgcc ctcgatgtat aacgatcccc agtttggcac | 4560 |
| | aagctgtgag atcactggct ttggaaaaga gaattctagt aagtgacaat tgcgactgac | 4620 |
| 25 | ttagaaggtc ctgaggagtg ttttgacctg aaaatgagcc cagtgtgatc aaggaagac | 4680 |
| | tgcagagtta gaggtgggag cactgaggcg gtggcagatg ggtccaggga tggatgaaga | 4740 |
| | gtgttgttta gggagcgatg ggctgcaaag gtaaatagat ggtaggggct ataggtggag | 4800 |
| | gtaaatggct cagatttga tggagagaga ataattggcc tctccctggg tgatgatact | 4860 |
| | ttatggtgtc ccctctctgg cgagacgtcc cacgtggagg cagataaatc ttgatgcaaa | 4920 |
| 30 | cgctccctg ttttctccac ctagccgact atctctatcc ggagcagctg aaaatgactg | 4980 |
| | ttgtgaagct gatctccac cgggagtgtc agcagcccca ctactacggc tctgaagtca | 5040 |
| | ccacaaaat gctgtgtgct gctgaccac agtggaagac agattcctgc caggtaggtg | 5100 |
| | ttccaagcat ctctctccac ctcttccata tctcccaga gctcctgggc ttgttccagc | 5160 |
| | cagcttaagg gtgtctctct ctagccaaag ccctaagtag ccagaatcag gagctcaggt | 5220 |
| 35 | ctttgagggt ttaaaccagt ccttatgtgt ttgccagaca ttacaaaaa aatcccagct | 5280 |
| | ctgcgctagt cacttcagac tgggggcacg agatcctaga aagaggaaac agtaaaagac | 5340 |

aatgtaactc agtgcacagg gtgtgtttgt aactataaat gatcaggtgt tcaggagagg 5400
gaggtgagtg ccaacctgag ggtcagggag gggaggtttt aaaggaaatg tgacttgata 5460
ggcatttgaa gaggcagagg gaagaaagga aggtgtttca gttgaaagat acaaaactga 5520
gaaggaggct ggcataattcc ggggtggggag gagaactagg gtctgggagt gtggatggaa 5580
5 tagtggcaga tgacagggtt tttaaagcca agcaggggat tttccaactt cgatgtggta 5640
gaaatggggc tgcgtcaggc acagtggctc atgcctgtaa tcccagcatt gggttaggcc 5700
gtagtcatg gatcattgag gccagagttg agaccggcct ggaccaacat ggtgaaaccc 5760
tgtgtctact aaaaaatgca aaaaaaaaaa ttagccagggt gtgggtgtgc ctgcctgtaa 5820
tcccagctaa tcaggaggct gagacatgga atcgcttgag cacaggaggc aagtttgacg 5880
10 tgagctgaga tcacgtcatt gcacgccagc ctgggcgaca gacgagatt ctgtcctccc 5940
gccgaaaaaa gaaagaaaat gggaagtcgc taaggacttt gactgggaaa ctcttcctc 6000
tctctggtat ggttgggtga tgggatcaga aatcccctcc tcacttctct agggctcatc 6060
ttttgtatct ttggcgtcac agggagactc agggggaccc ctctgtctgt ccctccaagg 6120
ccgcatgact ttgactggaa ttgtgagctg gggccgtgga tgtgccctga aggacaagcc 6180
15 aggcgtctac acgagagtct cacacttctt accctggatc cgcagtcaca ccaaggaaga 6240
gaatggcctg gccctctgag ggtccccagg gaggaacgg gcaccacccg ctttcttgct 6300
ggttgtcatt ttgacagtag agtcatctcc atcagctgta agaagagact gggaagatag 6360
gctctgcaca gatggatttg cctgtgccac ccaccagggt gaacgacaat agctttaccc 6420
tcaggcatag gcctgggtgc tggctgcca gaccctctg gccaggatgg aggggtggtc 6480
20 ctgactcaac atgttactga ccagcaactt gtctttttct ggactgaagc ctgcaggagt 6540
taaaaagggc agggcatctc ctgtgcatgg gtgaaggag agccagctcc cccgacggtg 6600
ggcatttgat agggccatgg ttgagaaatg aataatttcc caattaggaa gtgtaacagc 6660
tgaggtctct tgaggagct tagccaatgt gggagcagcg gtttggggag cagagacact 6720
aacgacttca gggcagggtc ctgatattcc atgaatgtat caggaaatat atatgtgtgt 6780
25 gtatgtttgc aacttgtgt gtgggtgtg agtgaagtg tgagtaagag ctggtgtctg 6840
attgttaagt ctaaatatct ccttaactg tgtggactgt gatgccacac agagtgtct 6900
ttctggagag gttataggtc actcctggg cctcttgggt ccccccacgtg acagtgcctg 6960
ggaatgtact tttctgcag catgacctgt gaccagcact gtctcagttt cactttcaca 7020
tagatgtccc tttcttggtc agttatccct tcttttagc ctagtctc caatcctcac 7080
30 tgggtggggg gaggaccact ccttactg aatatttata tttcactatt tttatttata 7140
tttttgaat tttaaataaa agtgatcaat aaaatgtgat ttttctgatg acaaactctc 7200
ctggtgcttg tatgggaagg agttggagta cataaaaagg agaaaataac aaagggtg 7258

<210> 56

35 <211> 852

<212> DNA

<213> Homo Sapiens

<400> 56

5 cagctgcgct ggaggctgag gccgattgct tgagcccagg atttgagggc cagcatgcgc 60
aacataatga gaccagctct ctaaatgcat gcctctctat atattaaaat tctgatgtga 120
aaatatttta aaatttaata catttcaa atgttttaatt gtataataaa caaatgtaa 180
ataataaaat aatttaatat taaattcaaa aatgaggtag aaacaaagca cagcgatata 240
aataataaat tttcctttac atttttgagg cggctctttg agttttggat ttccttctta 300
10 ggtcactgaa atgtgctcct tggagccagc ccgcaaatca cgcatttaga aaaacataac 360
tatacactcc taaccctaag tattagaagt gaaagtaatg gaatctcgat gtaaacacaa 420
tatcactttt ttgtagagct attttgagta taataaattt gaactgtgcc aatgctggga 480
gaaaaaattt aaaagaagaa cggagcgaac agtagcttcc tcgtccgctg actagaaaca 540
gtaggacgac actctccga ctggaggaga gcgcttgccg tcgcactcag ttggcgcccg 600
15 cctcctgct ttttctctag ccgcccttcc ctctttcttt cgcgctctag ccacccggga 660
aggcactgcg gtagctgggc tctgattggc tgctttgaaa gtctacgggc taccgattg 720
gtgaatccgg ggcccttttag cgcggtgagt ttgaaactgc tcgcacttgg cttcaaagct 780
ggctcttgga aattgagcgg agagcgacgc ggtgttgta gctcgctgcg gccgcgcgg 840
aataataagc cg 852

20

<210> 57

<211> 2501

<212> DNA

<213> Homo Sapiens

25

<400> 57

tcttgtcact ccatgcactg tgttccgtat gctaaatagt ttgagaaacc caaatgggcc 60
atgttcgcct acatttcatt gtcctgtact tcctgtcctg tactagcaaa gcagtcocat 120
30 tggctcttct tctcctcatt aacaataaag gtaacacttt tgatgttgtt tcttcagaaa 180
accttcattc atcaaaactg cctcaaagat catgtttgtt tgattccaga acttcctgta 240
attacctgtt attgtaacac tcatcactgt attttactta cttgtgtaac taattttcca 300
tattctgcac tagacaacaa agtcctttaa gtcaggtagt atatctatct acatagcatt 360
cacatctcct acaataaggg acattagcag ataaacaaca catattaaat gaataatgaa 420
35 gtttctgaaa tactacagtt gaaaactata ggagctacat tatatagaat aaacatttac 480
tttgcctatag aattcagttg aaccagga ttattttatc ctcaagtctt aggttggttg 540

| | | |
|----|--|------|
| | gagaaagata acaaaaagaa acatgattgt gcagaaacag acaaaccctt ttggaaagca | 600 |
| | tttgaaaatg gcattccccc tccacagtgt gttcacagtg tgggcaaatt cactgctctg | 660 |
| | tcgtactttc tgaaaatgaa gaactgttac accaagggtga attatttata aattatgtac | 720 |
| | ttgcccagaa gcgaacagac ttttactatc ataagaaccc ttccttggtg ctctttatct | 780 |
| 5 | acagaatcca agacctttca agaaaggtct tggattcttt tcttcaggac actaggacat | 840 |
| | aaagccacct ttttatgatt tgttgaaatt tctcactcca tcccttttgc tagtgatcat | 900 |
| | gggtcctcag aggtcagact tgggtgcctt ggataaagag catgaagcaa cagtggctga | 960 |
| | accagagtgt gaaccagat gctctttcca ctaagcatac aactttccat tagataaac | 1020 |
| | ctccctccca cccaaccaa gcagctccag tgcaccactt tctggagcat aaacatacct | 1080 |
| 10 | taactttaca acttgagtgt ccttgaatac tgttctatc tggaaatgtg tgttctctt | 1140 |
| | catcttctc tattgaagcc ctctattcc tcaatgcctt gctccaactg cctttggaag | 1200 |
| | attctgctct tatgcctcca ctggaattaa tgtcttagta ccacttgtct attctgctat | 1260 |
| | atagtcagtc cttacattgc tttcttctc tgatagacca aactctttaa ggacaagtac | 1320 |
| | ctagtcttat ctatttctag atccccaca ttactcagaa agttactcca taaatgtttg | 1380 |
| 15 | tggaactgat ttctatgtga agcacatgt ccccttccact ctgttaacat gcattagaaa | 1440 |
| | actaaatctt ttgaaaagt gttagtatgc ccctaagagc agtaacagtt cctagaaact | 1500 |
| | ctctaaaatg cttagaaaaa gatttatttt aaattacctc cccaataaaa tgattggctg | 1560 |
| | gcttatcttc accatcatga tagcatctgt aattaactga aaaaaataa ttatgccatt | 1620 |
| | aaaagaaaat catccatgat cttgttctaa cacctgccac tctagtacta tatctgtcac | 1680 |
| 20 | atggtactat gataaagtta tctagaaata aaaaagcata caattgataa ttcaccaa | 1740 |
| | tgtggagctt cagtatttta aatgtatatt aaaattaaat tattttaaag atcaaagaaa | 1800 |
| | actttcgtca tactccgtat ttgataagga acaaatagga agtgtgatga ctcaggtttg | 1860 |
| | ccctgagggg atgggccatc agttgcaaat cgtggaattt cctctgacat aatgaaaga | 1920 |
| | tgagggtgca taagttctct agtaggtgta tgatataaaa agccaccgga gcaactcata | 1980 |
| 25 | aggcacaac tttcagagac agcagagcac acaagcttct aggacaagag ccaggaagaa | 2040 |
| | accaccgga ggaaccatct cactgtgtgt aaacatgact tccaagctgg ccgtggctct | 2100 |
| | cttggcagcc ttcctgattt ctgcagctct gtgtgaaggt aagcacatct ttctgacct | 2160 |
| | cagcgttttc ctatgtctaa atgtgatcct tagatagcaa agctattctt gatgctttg | 2220 |
| | taacaaacat cctttttatt cagaaacaga atataatctt agcagtcaat taatgttaaa | 2280 |
| 30 | ttgaagattt agaaaaaact atatataaca cttaggaaag tataaagttt gatcaatata | 2340 |
| | gatattctgc ttttataatt tataccatgt agcatgcata tatttaacgt aaataagtaa | 2400 |
| | tttatagtat gtcctattga gaaccacggt tacctatatt atgtattaat attgagttga | 2460 |
| | gcaaggtaac tcagacaatt ccactccttg tagtatttca t | 2501 |
| 35 | <210> 58 | |
| | <211> 2501 | |

<212> DNA

<213> Homo Sapiens

<400> 58

5

10

15

20

25

30

35

attaattctg caaatttta taaatgcttt attttaagct aaatgctgag atgaaaaaat 60
gaaaccatat gagttagcaa agtagaaaat ataggcatat taatcagtaa atgcagaatg 120
ataaatgctc catcaatatg cacttggtgt agtgaggcca ccgaggagg tgcaatcctc 180
tcaacctggg aggagcaggt aggacttcag atgtcatcca actcaaagat atagtggagg 240
acttgatcaa acatttgcca agaccactat gagttaaatg aatagattag gcatttctcc 300
aatgttgcaa gcttcgaatc atatccaaac tcagaacaac atagcttggc cataatgatc 360
ccaaggatcc tattggccat tgtctttgag cctcaaagga acataataaa actccataat 420
acccttttga tctattctga agttaagtag tgaatttaca tgatgatgac acaaactg 480
taaaggacct ctgggttact tgtttataag ctagtatttc ctgaatcaat ttttctgac 540
cctagatatt tggtaggtga agtcatacct atatatcccc acaccctaga acagcatctc 600
caacttattt ttccctcctt gtcttttagt gggagccaca tcagtatcca agaggagatc 660
cagaagcctc tccaaccagg tagggacagt tatagattcc agacctcagc tatggccttt 720
gttacagagt acaaatgtta tatagtacaa gtttattgta cacatcccat tgagtctctg 780
agcttttagaa ttttcttgta gaatttaaca gttttttcat gccgtattta catattattg 840
ctagtattta gaattttctt ctccaaatgt ataacgttta ttattgcatt ttttgtatcc 900
actaagtgga aaatcatgca ttagatattg tagaagtaga tacaacaatg aacaagaact 960
ggctctgacc atgagaggaa ctgatgatcc aatgggggag atagacctgc acgtgtttaa 1020
taaaaggaag tggctattcc ggtttctttt tgatgggcaa gcattttgca aggccttggg 1080
ctatgtgtgt gcaaggctaa gccagttagt taattgggat ttttttaaaa aggcacttca 1140
ctggggggaa aaggaacata gagttgggta ttgtccctt gcctataata aaaacctatt 1200
atttttaatt ttttaactgg gtttgcggtt aaatctcaca gcccaagaga tttgccactt 1260
cagatggatt ccatacactt gcatttaagt atgcaaaaaa attccaatta tccagcaatt 1320
taaccaaatt attgtaact tttctaaaac aaaaaaaat tgtttccctt gttttggcag 1380
caatttcagt tacagtcctt tactttctac tcaagaaaat agtttcaaaa agttgatgtt 1440
tgttgctaaa agaactattt ttatgaataa atataaaact aagaagttat ggtgtccctt 1500
ttttaaaaaa tgactcatca aaagaaataa ctttttccct tctcttgtaa gagaaaaaa 1560
ttaatctctt ttagaattgc aaacatattt ccttgatgga gaaaatcaat tcacatggca 1620
tagtcgttat ttatccagtt caaaaaccag agtagaattt actactctgt ctccattttt 1680
tctctccca ccccttaac ccacattgga ttcagaaagc ttcattctgc aatcagcatt 1740
gtcctttatc tttccagtaa agatagcctt ttggagtcga agatgaggaa aagcctgtat 1800
tttatagtct tggaagtgtc ttcttttgcc aggacagaga gaggagcttc agcagtgaga 1860

gcaactgaag gggttaatag tggaacttgg ctgggtgtct gttaaacttt tttccctggc 1920
 tctgccctgg gtttccctt gaagggatth cctccgcct ctgcaacaag accctttata 1980
 aagcacagac tttctatttc actccgcggt atctgcatcg gccctcactg gcttcaggag 2040
 ctgaataccc tcccaggcac acacaggtgg gacacaaata agggttttgg aaccactatt 2100
 5 tttctatcac gacagcaact taaaatgcct gggaagatgg tcgtgatcct tggagcctca 2160
 aatatacttt ggataatggt tgcagcttgt aagttatttc ctttcatctg tttcaaagt 2220
 tagcattcaa ttttagccct ggttttggct tcagtcagtt ttgcgatagt agtgaagtaa 2280
 agacactagg attttaacaa gtaggaaaag ttaatttagt ctaactttta atatgcaatt 2340
 gagttttgct atataccatt gtactgtcat agttagagct gaaaattgat gtttttggt 2400
 10 tctttttttc caaaggcaat tgagtaattt ggattctgtc tctagtcggt ctgtctcttt 2460
 agtttcctat acttgacaat gaggtcaaac ttagcaaata a 2501

<210> 59

<211> 2501

15 <212> DNA

<213> Homo Sapiens

<400> 59

20 ataaaaaag acatgaaatg aatcggggaa aatatttgct acataactaa gaatgaaggc 60
 ccttaataaa atctgtaaaa ctatacacac ttttaggaat gaatcaacaa ataatttcta 120
 tgaattagaa aaaagtgaca atccaactaa aaaatgaata agggatataa gcaatgtgtt 180
 tcacagaaaa aataaaaatt gacaatgaag ttatgaaaaa atgttcagtc tccttagtaa 240
 ttgcacaaaa caaactaaaa caatgagaca ttacccttaa gattagtaaa tgttaaagaa 300
 25 aaataataat tggtagagggt gtgggggaagt gggcacttac acctatgttt ggaaatataa 360
 attggtgcaa ctttataggg agagcaatct cacaacattt tccaaagact tacatgcaca 420
 accctatggc agagaaattt attcctcttc caggattttt tttccttcaa aaacagtgat 480
 gtggatgaaa aacacatggt cactactgca cagggtataa cagctgaaaa ctggaacga 540
 taatactcac attcccttca gtaggggaat ggttaaataa attttacaag ccatctggta 600
 30 gataccaggc atgagctaaa agttagggtc cagttagaga tggaaagcac accagtaatt 660
 tgaaagggaa aatgtaatat gaagaattat taactagtaa aagaaggcta actgctaaag 720
 gtacaagagc actcaagctg tctgcagtca gcaggccccg gctggtgagc aggaagctgc 780
 ccgctgggag gctgccaaag ttccctgaag gtgagcacca ctggttctac aagctgctgg 840
 cagtcatggc gttaagagca ggaagagaag caccagaacc cggaagagaa atccagtcct 900
 35 ctgctaggcc ttgcaccgtc cctctggcgc cctctactga caaagccagt aaaattgtgc 960
 cgctagcaaa ggagatcttt ttatgggatg tagcttggtg tcaccaaaga gaacagagt 1020

gacttggagc tcagatgcaa cacaatgatt gatactggca cagtatactt accctgcttt 1080
 tgtaaacaaa atggtatatg tgatgtctct ctttgtctct ctgtatataa aacaatattt 1140
 gtttctactt attatgtatt tatgtcttta ctctgcatgc caggagctaa gtattttgca 1200
 tgtattaact cttttgttc tcataataac cttcacatgc aggaatcatt atagctactt 1260
 5 tatgaatgag ccgaggaagg cactgagacg ttaagtaact tgcccaaggc cagcagcta 1320
 gtaagtggca gagcaagaat tactatggct ttataagcct aggaaaaagt ctgaaagaat 1380
 caaaatgtta acagcgggga cctcaaggaa gcattgaaga ggccatggga gaagttttca 1440
 ctttgttaaa aaatcagtcc ttcaataaa taaatacagt gaggcttccc cagaagcaga 1500
 tgtcactatg cttcctgtac agcctgtgga actgtgagcc agttaaacct cttttcttta 1560
 10 taaattatcc agtcttaggt atttctttat aacagtgcta ggatgagctg atacagtttc 1620
 ctacactgta acctaaggca atgctttgca caaagggatg agccagattg cttagtaatt 1680
 aaaacgcaaa tacaaaccac aagcatatcc attcatgaat tggggggctg ctttgtgtgc 1740
 atagataagg tatatttttt aaaaaatta tttttccaag aagaaaataa accagttaat 1800
 aaacgacaac tcacagtgcc aggaagtgaag aaacaagtgt gtgataaacg gtggagaatg 1860
 15 ggagcactct ccgcagtggg cgggaggaga cgaggagggc gttccctggg gagtggcagt 1920
 ggttggagca aaggtttgga ggaggttaagt catgtgctct gattttttgg tttctgtttc 1980
 acctgtgtc tgagctggtc tgaaggctgg ttgttcagac tgagcttcct gcctgcctgt 2040
 accccgcaaa cagcttcaga agaaggtagc tgggtggctgc ctgaggaata ccagtgggca 2100
 agagaattag ctttctgga gcatctgctg tctgtgagat taagcactat gtatattgct 2160
 20 ttattcactc cccacagcaa ccttaccagg cagttctttt ccacgtgaaa agatggaggc 2220
 tgggtggagc aaaaggaggt atttagagtc ctcagcaagt gagaggcaga gctgggattt 2280
 gaatccagat ctgcctgata ctgaagtcta ggctggttcc acctctccgg actgctttcc 2340
 agggagtaga agacagatat ttaccttag ctggtgctt ctagaagtct gacctgctg 2400
 gctcaaaacg actttagtgc cttgccaga ggctgcgggc tgcgggtcaa gacatcagta 2460
 25 gaaggagggc ccagccagag aggtgacat gggcttctac t 2501

<210> 60

<211> 2501

<212> DNA

30 <213> Homo Sapiens

<400> 60

cgggcaggaa taatcactgc ctccatccc cttaaacatg ccaagatgct ttatccctag 60
 35 gatgaggtga cttactccag gtaactccta ttgcctaacc actgaccaat tactctgccc 120
 tttagtcttt atgtcattaa atctgcatta agaatttcat ggaataggcc cggcatgggtg 180

| | | |
|----|--|------|
| | gctcatgcct gtaatcccag caccttggga gaccgaggtg ggaggatcac ttgaggtcag | 240 |
| | cagttcgaga ccagcctgga caacatggcg aaaccccatc tctactaaaa acacaaaata | 300 |
| | actagccagg tgtggtggtg ggcacctgta atcccagcta tttgggaagc tgaggcagca | 360 |
| | ggagaatcgc ttgaactggg gaggcagagg ttgcagtgag tcgagatcgt gccagtgcac | 420 |
| 5 | tccagcctgg gcgacagagc gagactctgt ctcaaaaaa aaaaaaaaa aaactcaggg | 480 |
| | aatggatagc agcattgatg aatattgcgt ctggagagat cagatcactt gtcacttgtt | 540 |
| | tccaggcaca gggcttacca agaggcagat tccagattta aataattctg taacagcaaa | 600 |
| | gtccaagcta ttttactgc tttggagaaa agaccagac ccagagcttg aacctcactt | 660 |
| | tgcagacccc cagttctaata cttttaagtt tttttttttt tttttttttt tttctgctgg | 720 |
| 10 | gcacgggtgt tcatgcctat aatcccagca ctttgggaag ccgaggggga aggatcgctt | 780 |
| | gaggccagga gttcgaaacc agtctgggca acatggcaaa accccatctc tacaaaaaat | 840 |
| | acaaaaatta ggccagagtg gtggcgcgca cctgtagtgc cagctacgtg agaggcggag | 900 |
| | gtgggagaat cgcttgaacc cgggaggcag aggttgcaat gagctcagat cccgccactg | 960 |
| | cactccaggt tgggcgacag agcgataccc tgtgtgaaac tttttttttt ttctccaacg | 1020 |
| 15 | ggctttccag agaagtgtgt gtatgtgcgt gtgtgtgcgc gagcgtgctt gcttgggctt | 1080 |
| | aaactttctg tcgggccaca ctttcccaag tctttgcact ggctgtaggg tgggctttat | 1140 |
| | cctcgggacg tcctcctccc caagtccagc ctgcagctgg aagtcttcac tgatctccat | 1200 |
| | ctctcctccc tgatctcctg ctctcctccc tgccgcctc aggactggga ggccgatctc | 1260 |
| | tctctctcgc cctcccctcc accagccttt tccagatgta tgtctgcaa agacccccca | 1320 |
| 20 | gtgcagagga tgatgaatga agatcctcga gccagcccg tgggaaagt tctctgccta | 1380 |
| | caaaagcgag ggaagggaa ggaagtgtg gggtagggga aaagttagag ctgagaggct | 1440 |
| | ggggcgcgac gagtctggac accgggcggg gaccaagct ctctccgctc agccaataac | 1500 |
| | tgtgcctccc ttaggaaggc gtgaggaaat gctccaatca atccctgcac tcctcccttg | 1560 |
| | gaatttgggc tgtatttttt tatttactgc aaaccacaca atccaccag gggtttcccc | 1620 |
| 25 | agtgtttgcc tccagcggtc ccggtgccca ttactagtg ctgctccctc tcttccgcaa | 1680 |
| | gactgcgctc cagtcccagc ctcttctcc gcgggtgcct cccaaaccgt tctatcattc | 1740 |
| | tcgggttcag ggaggcgga tctgtcctgc tctccggttc ctttaagagg cgtcggctcc | 1800 |
| | accctctca gagtccgggt ctgacgcgag atgacagcaa cgagtccgt atgtctatgc | 1860 |
| | aaataagcgc cctcttgtgg gccaatgggg agcggagggt ccggaaccac ggaccaatgg | 1920 |
| 30 | ggcggggcg ctggggctca ccatataagg agcggcctcg ccataaaagg aaacattgta | 1980 |
| | tctctttata tggggggaag ggtcggggga tccctccgcc gccagcgcgt ggtcccggcc | 2040 |
| | ccctccaccc gccgtctcgg ccgcggccag cagcccctgc cccccgggg acgctgacgg | 2100 |
| | ccgcccggcg cgccgcccta gcagacggac agggggcgct gcgcgcggcc tggggcaacc | 2160 |
| | cgggccacag gggcaggaaa gtgagggccc aggtcgcccc gggcgtgcag gggccccggg | 2220 |
| 35 | ttcgcagcgg cggccgcggc agcgatagcg gactagcag cagcgggagt gccgggttga | 2280 |
| | gccgggaagc cgatggcggc ggctgcggcg gctccgattc ctgctgact gccgctccgc | 2340 |

| | | |
|----|--|------|
| | cctcctgcat cgagcgccat gttaccgacc caagctgggg ccgcggcggc tctgggcccg | 2400 |
| | ggctcggccc tggggggcag cctgaaccgg accccgacgg ggcggcggg cgcggcggc | 2460 |
| | gygacacgcg gggtacacgg ggccggggtc cccgggaatg g | 2501 |
| 5 | <210> 61 | |
| | <211> 2501 | |
| | <212> DNA | |
| | <213> Homo Sapiens | |
| 10 | <400> 61 | |
| | ggaaccctct gatagagagg gctgactgta tttattgaaa acaaaacaaa acaaaacaag | 60 |
| | ggttggtattg gtggacccat gcagctcaaa cccttggtgtg tcccagggtca actgtatatc | 120 |
| | cagagcttat aggaaaatac ctctcccagt aaccctgtc accattttctc tcttaagcta | 180 |
| 15 | ttattatgat tagccacggt ttgctattta aatttaaatt taaataaaaa tgtggccttt | 240 |
| | cagttatgct agccacatct aaagtgtca atagccatat gtggctaata gttactatct | 300 |
| | cggacagcac atatttagaa cattcccatc atttcagaaa ttttcattgg gaacactctg | 360 |
| | cggaaaaagg gggccatcat aatgtgagtc catcttcttg aaaaatcctg ggaaggggac | 420 |
| | aaaggaggtc tgtttggcat tgtgtaatgg taatttggtt ttttaattttc aaaaatgttt | 480 |
| 20 | acccaattcc tattcatcag ccagggtgtg tggtctctgc ctgtaatccc agcactctgg | 540 |
| | gaggccgagg tgggaggact gctgcagccc aggagtttga gaccagcctg ggtaataata | 600 |
| | gggagatcct gtttctacaa aacacacaaa acaaaacaac aactttgatg ttgtggagtc | 660 |
| | aggacagtcc tgggttaaaa cctttgctct ccttagctgt gtaaaccgtg ggtctcagct | 720 |
| | ttcttatctg ttaacggtag gtacttcttc ctagggtgtg tttgaggatt aagtgaaggt | 780 |
| 25 | ccaagattgt gtctggcaca cagtagcttc tcagcaaatg ttttcctcct atgtcagggg | 840 |
| | atggctcctt tatcccgttt tgggcccatt ggtggccctg aagggtgggt gctcaggtgt | 900 |
| | taagttctgt agatggcata tccttgggaa aagcaaggca attaaaaaca gtgagaggtt | 960 |
| | gctctgggta agttttctcc tataactttc cccatgggtc aattgggtag aatctgccat | 1020 |
| | tttctaataa cttactgatg gtagtggcat tcggaagcac aatagctgaa gccggagctc | 1080 |
| 30 | tgagtggaga gaaaggctctg tttctcaggc ccaaaaagag gttacacacc catggctgtc | 1140 |
| | cagtttggtg gtgcaggccc tgaaatcaga ccaaactgga tttaaatccc caaacctata | 1200 |
| | ctctaagcta tgtgacctg ggctagatac ttcacctctc tggccttatg aagtaggaat | 1260 |
| | aataataata ccgtctaggt tgtaggagt attaatgag gtaaagcact gaaaacgttt | 1320 |
| | agggactgtg ttaaatcatt aaataaataa aaacgggat gaccttatcg gcttgacaca | 1380 |
| 35 | ggggattaaa tgagataata tatgaagaca agtacacggc aaatgcttaa ttaatgttgc | 1440 |
| | ttatttttat gtctgcaaac tgacttaaag gggaggcctt taagaaagac agtggggcaa | 1500 |

5 tttgcgcgtt gatgcattgt aggagaaaat gtgcaggggg cccgttggga ccagagttca 1560
 accaggttaag cggcagaaaa ccacaaatac ctccaggcgt tcctggggca gcgcgcctc 1620
 cccaaaatca cgcaaaactt ggtttgctaa gaattgtcag ctcttctaaa ggaggcgctt 1680
 cacgcattctc agtctgtgaa atgggaccca ggaccagggt agagggtcgt tctcggcctg 1740
 gggaccgagt attttgtgcg ctccggtaac gcaggaagac agcgccactg acactctaga 1800
 gaccagcggg caccgcctgg aggcgccttc accacttggc ggttccgggt ccgcgcccc 1860
 ccgcgccaca agactcacgc ccgaaccacg tgatcagggc cgtggctccg ccccgctccc 1920
 gcgcgcgcgc ccgcttccgg tagggcgga aagcggaagt gtgggagggt ctgcggggcg 1980
 ggctcaggag gtccgcggga ggatggagca gtgagcgggt ctggcgcgct gctggcagcg 2040
 10 ccatggagac ggtacagctg aggaaccgc gcgcgcgtg aggggccact ggctaagagg 2100
 acgggcatgg ggtcagggga agaaaaggcg ggaactggtt gaggggatac acctgtgtgg 2160
 gagtccccg agctaagcga ccagccgat ggggcacctg ctgagtgagg ggggggacgt 2220
 ctggtgggtg agggtcggc tgaggggagc atctgctaag gaggttagac ttgggaccg 2280
 ttagaggag cactcgtgt ggtgagactg tgctgaggaa cgtggggaca agttaggag 2340
 15 agtacctgct gagggcgggc cactcggggg aacgtatcc aagcaggac tcacggaggt 2400
 gggggcgaat gctgaagcag ggtgagaatc tgtgaggat ctctttaagg ggggtgatcg 2460
 agaactggcc aagaggaagg ccgggtggac tttctaagg t 2501

20 <210> 62
 <211> 2501
 <212> DNA
 <213> Homo Sapiens

25 <400> 62
 gcatggtggc tcacgcctgt aatcccagca ttttgggagg ccaaggcagg cagatcacga 60
 ggtcaggaga tcgagaccat cctggcgaac acggtgaaac cccgtctcta ctaaaaatac 120
 aaaaaattag ccgggcatgg tggcggggcg ctatagtccc agctactcgg gaggctgagg 180
 caggagaatg gcgtgagccc aggaggcaga gcttgcggtg agctgagatg atcgggccac 240
 30 tgtactccag cctgggcaac agagtgaggc tccgtctcaa aaaaaaaaaa aattactaca 300
 tgatactaag taatgcggaa ggtgactcaa agggggaaag gaacacagca gtgtaaagga 360
 aggaggttgt agatggatct agaatttccc cctcatttcc atcaggtgaa agcctgagaa 420
 aactgcaatc tttgtgcagg ctgggtttgc tttgtacaca ctggtccctt agtgttcac 480
 tccaataatg ctgacaactc tgaaaaccat ctgtagacat tctgcaggct ccattctcagg 540
 35 aacaatggct attttttcgg gtagttgaag caaaattaag tccaatgata agcaaatata 600
 accattatca aaatcttcca tttatgtttg ttaaagcaac ctaagtatga tctgagaagg 660

| | | |
|----|--|------|
| | actctgtatt ctatatttga gtccttgtgg atgaactgta acctagctta ataggcagac | 720 |
| | aagattgaaa acctaattha ggagtatgtg cctttaacaa tagctgagtc ttggccaatc | 780 |
| | ccagtggcca tacttcaacc attcatacac tgctgagtgt tcaaactgtg ttcaaagaag | 840 |
| | gcaaaagcca acctgtaacc aatccagttg tttctctgcc ttacctcaa tttctgtatg | 900 |
| 5 | tcacttccct ttttttgtct ataaatatgt tctgaccatg aggcatccct ggagtctctg | 960 |
| | aatccgctgt gattctggaa gctgccccat tcgcaaatca ttcattactc aattaaactg | 1020 |
| | ctttaaatth aattctgctg aagttttctt ttaacagggt tagaaaaaat aatggcaaaa | 1080 |
| | atgaatgaaa atccaataac cctggaagca gaaaaggctg ggggctcaa taagtgtaaa | 1140 |
| | tagtcccac cctatatttt ctccatggca attacaatcc agcacattat atatatttt | 1200 |
| 10 | ttttgcttct cgcatttttg cttagggtaa agctttttta aacaggcact gccaccagt | 1260 |
| | gttatcaaga aggtctggat gccgttttgt gggaacattt taaagaggaa tgtccaaaag | 1320 |
| | gaaaaggggg atgggttggg agaagggtat caggcgggta tctcaaaacc attcttaggg | 1380 |
| | ctataggttt aatttatttg gttgtggacg tcagagccgt catggttaaga aggaagcaaa | 1440 |
| | gccttttgta ataattaaag ccttcagaag cagcgtgccc cattgcccac tagtgcgccg | 1500 |
| 15 | tgaagtctg tgttcaccta cagggtccct ctgagcactg cccaggcctc ccgagtgtc | 1560 |
| | cagcacagta gcttgagct tggttggttg gtgaccaaga tacactccag ggaatatgcc | 1620 |
| | atgcagtgga gtctcttccc cggcactgca tagcaaaagg aaagggccgc tgggtgtctg | 1680 |
| | tgggtcctgg gcagtcacag aagccaccgc gctggcggg aggaggggga ccgatgcggt | 1740 |
| | ccatgtcccc ggcagcccca ctttctctgc ctgcgaagg ccctgttccg gcgggaggag | 1800 |
| 20 | agaggcgcgc cccaccggg ctctctaca cctgccgccc cctgggcca ttccgcgggc | 1860 |
| | ctgcgccgc gcttcagccg attcccgcgc agctccgggc tcatgggcgc ggtcagcagg | 1920 |
| | gcgggccagg gcggcggggc gcgacactgg gaggaagtgc gggccgcctg cccgggcgcg | 1980 |
| | ttaaggaagt tgcccaaaat gaggaagagc cgcgggcccg gcggctgagg ccaccccggc | 2040 |
| | ggcggctgga gagcgaggag gagcgggtgg ccccgcgctg cgcgcgcct cgcctcact | 2100 |
| 25 | ggcgcaggta ggtgtggccg cgtcccctac ccggccggga ctttctggta aggagaggag | 2160 |
| | gttacgggga acgacgcgct gctttcatgc ctttcttctg tctacattca tcggccgagg | 2220 |
| | taaaagtgt gaaacatgt gaataaaata cagggtgggt cgcgcagctt cgtcctgaa | 2280 |
| | cctaccgcg ctccgggatcc agaagctgcg ccgggagaga ggggctcagg cctgggcgga | 2340 |
| | ggggacggag gtcagaccgt gcggaagtg acccgggcac cccagggcgc ccaggcccc | 2400 |
| 30 | agggagcgcg gaaagtgcg tcgcggcccg gccctcgga gacgcgggat tgggatcagg | 2460 |
| | cacagcgcga ggaagtcgat cttggagcta gaacattttc c | 2501 |

<210> 63

<211> 2501

35

<212> DNA

<213> Homo Sapiens

<400> 63

| | | |
|----|---|------|
| | cccaaaagat acaaaggggt ataaggtgaa aaattattct aacctatccc tcagtgcacct | 60 |
| 5 | agttcccttc ctctgagggtg accaattttct tgtgtatctt tcctgagata atctatacat | 120 |
| | atagcaccat atacaagcaa atgaaatatg ttttatttat ttttttgaga ctgggtctca | 180 |
| | ctctatcacc caggctggag tgcagtgaca ccatcttggc tctccgcaac ctctgcctcc | 240 |
| | tgggctcagg tgatcctccc accttaacct ccagagtagc tgggactaca cgctcacacc | 300 |
| | accacaccca cctaattttt gtttttttgt agagacgggg tttcaccatg ttgccaggc | 360 |
| 10 | tggtctcaaa ctctgagtt caagtgatct gccacctcg gcctcccaa gtgctgagat | 420 |
| | tacaggcgtg agcctccacg cccggcccca aaatctgttt taaaagcaga catttcttgg | 480 |
| | tgattctaataaagggggtt ctacagacata tttggaaaaa tatatoccta cttttatgcc | 540 |
| | agaccctgtg ctgggtcccc gggctgtgtg acctgacact gcacagtcct gcttagaatg | 600 |
| | cttaaagaga gttaataagg taccaccttc tatgccatag gcggggagca aaggggctcc | 660 |
| 15 | agtgggcctt gcctaggagg cctgaagcta gagctgctga gggcagggt gtgctgcaaa | 720 |
| | gaaaatgtct gagagctgca ggcgtttcat cttctgtcat cagctgtggc acctggcaga | 780 |
| | cactggatag gcttgttagac aaagacctgg taactcaagg agctgcttgg ccttcctgcc | 840 |
| | cagtcctc cagaggcac tgtacatctc tggtttcttc agggggccct gtgtggaagt | 900 |
| | atcttttgtc ttcttggtgt cagggatatc atcacgtgcc tgttggttag gcgagcccg | 960 |
| 20 | cgccagctct cctaggatgg ggagagtaat gttcccgagc agaacagggt ggggctttca | 1020 |
| | gactactccc tttcctttac agctggcttc attccatcga cctcatcaaa gccttcctgg | 1080 |
| | gagcacccta gagaagagtt acgtccaggc cgggccctgg ctgcctggtt caccggcgaa | 1140 |
| | tccccagcac cagcctcgc acgtcgggt caaagcatgt ttagtgaagg agtaggtacc | 1200 |
| | tactgctaga tggagccatc tctctagact tggggtttcc ctataacgat ggctatgttt | 1260 |
| 25 | ggcatggaag cctctttaga agtcaatagt aggaaataag ggctaacagc acctaatgt | 1320 |
| | ggagtaaggt tcaaatccta gctctgccac ttaaccgttc cgaacctgtt ccctcactgc | 1380 |
| | agaggcgaaa aggctaacac tatttcacct cggagggtta ccgtggagaa tggaagctgg | 1440 |
| | acaagctgta tcagttcagt agtaaaacac acacacacaa gcgccccacc cccacccac | 1500 |
| | cccacccag gaatgaacac acacaccgc gcgcgcacat acacctcagg aatgaacaca | 1560 |
| 30 | cgcgctaca cacacagca gccccccca ggagtgaaca cacacacaca cgccccgttc | 1620 |
| | tgttgttccc aggaacacac acagagacgc acacactcgc ccggttttgt tttttccagg | 1680 |
| | ctttttaact ggggtcttcc actcggctta gggcacgct gcctgaaaga cttttctagg | 1740 |
| | ccagtcgggg tccggcacc agttgacgag acagcgcggc gctttcagag ctggggagag | 1800 |
| | gcgaaaactc ttccggcccc ccgatcccc gccagccgc ccccggcage tcttgccgc | 1860 |
| 35 | ctcccgccct gggcccgccc agccgttctc ggctgccgt caggcgatct cggcgccag | 1920 |
| | cccagccgcg atgtgacgcc gcgcgcccc gggctcctcg gcctgcgc ctctcctata | 1980 |

| | | |
|----|---|------|
| | aagcagacgc cgcgccgcgc tgcgacgctg tagtggcttc gtcttcgggt tttctcttcc | 2040 |
| | ttcgctaacg cctccccggt ctcgtcagcc tccccgcggc cgtctcctta acaccgaaca | 2100 |
| | ccgtgagtag ccgcccactg aactggaaaag ggtcgtggct accggattgc gtgccggctg | 2160 |
| | gcctcaccgc tgcggtttgg gcctgcccgc ggcggggcgg gactgggcct ggccttcttt | 2220 |
| 5 | cgggccccgg ggatcgcggt gtcgaccctg ttcttcggga gacactacca gggtccgttc | 2280 |
| | acctgccccg cccccgactc agcgaggcct cctctggcgg ggcgtcctca cggcgctcca | 2340 |
| | taagtgaagc gaacccccgg ctgggccttc tctgcaccgg ccgagcgtca gccggcgcgg | 2400 |
| | agctcggctg caaggcccag gctcgggcgg ggggcctctc ttggtcttaa gcctgtgtgc | 2460 |
| | ccggggacca gggcgggggt ggcggcgggg ttgtgaatgg g | 2501 |
| 10 | <210> 64 | |
| | <211> 2501 | |
| | <212> DNA | |
| | <213> Homo Sapiens | |
| 15 | <400> 64 | |
| | | |
| | gatctgacag gttaaagggtg tacacttatt ttctctgtaa gaagcgtcat ctggtgaagat | 60 |
| | gatcaagaat ggtgcaaagc aggatgggga gtttaaaatt gtttccaaat gtgggaatgt | 120 |
| 20 | aatgaatat aaacatgtaa gattttaata taccaaactg atcagattct gtgtaatttc | 180 |
| | caagtttctt ttttctttca aaactcctct gaaatctgac tgtccacaaa aacttacttt | 240 |
| | atagaatttt atgtgattta ttactcaga tattatactg acctcacatc cagtagtgaa | 300 |
| | aacagatttt attgtagaat ctggaaagat agagggccat ataggttgta ttttcagttt | 360 |
| | tgtttatact aacacgtgtt tacaaccag tttaatctac acctgtatt gtattattgt | 420 |
| 25 | tgtcatatct ctgtatgcat gtaagtataa tatgtgttgg caaaggaaaa ttttgagtaa | 480 |
| | gaagaagctc tctgatctat ttgattcaat atgtatttga gtgtctaaca gacactgttt | 540 |
| | tagacactgg tgatacaaca ctgaacggag caccaaatac ttacagcgt ctctggagc | 600 |
| | tgttgtcaag acatactttc caaggggaat atttcagaat aggtgataac tagtcaacga | 660 |
| | aggaaaagta ccttagtcat ctaggagagt tgtacttaga gtgaactgaa ataaactaag | 720 |
| 30 | ctcacgaaag acagagattt ttgtttggc tttgtctgt tgcatctact actgtatctc | 780 |
| | cagggcccaa aatagtgtc ggctcataat aagtattcag caaatatatg ttgttgattg | 840 |
| | gagtgtttgt tttgaatttc tgtaatcaaa cacatacctt ggtaaattat ctttacatct | 900 |
| | tgctagttag aaattttatc tcagttgctt tgtttttaat gttacottgc tttttgttc | 960 |
| | tacttgtgcc atacatcagg atgtggaaa agcttattaa tattgacagt catatggtta | 1020 |
| 35 | tctgatattg aaaagaatag atttggaaag gaacctaaag ggtcatcttt tgttcagctt | 1080 |
| | cctgcctagg aaaactaagt aagatgatta ggtatgtata ttttaattagt catttaaaaa | 1140 |

| | | |
|----|--|------|
| | aaaaccagga caacataatt gagttccctc ttgagaaaat ggagaaaggt acttaaccct | 1200 |
| | agctataaag ggactaacct ggaaatttta gaacttctgt gtgggaaagt ggaaaaaaaa | 1260 |
| | aaaaagcaca actaagctgc tctttgttga tatcagaaat gggcctgtca ttcatttttg | 1320 |
| | cattgaagca tagcctccta tctcggggca ggactgggac atttttttcc tcccacaaga | 1380 |
| 5 | gctggacagt tattacaggt tcaaaaagcc ccgaccagtt tttcaagagt ttctcctcct | 1440 |
| | cttttcccc tgaaactcgt ggtgcttttg ctctgctttc aagatgcatt aagtctcctg | 1500 |
| | ctttgtgact gctttggagc cagcagatac tctgatatgt ataattcaaa ttatgcaggt | 1560 |
| | ttcacgagta agtttaatct tattttttaa gttagttaaa aggcaagtga tatttagaaa | 1620 |
| | aatgttaact tgtagttatt tcaccctttt tactttaagc atttttattg cttctcggcc | 1680 |
| 10 | ttttggctaa gatcaagtggt gtactttaag cattttttta aataaaaata tccttttaat | 1740 |
| | ttaataagaa aacaagggtc tacatagaaa agccccctca tctaagacct gcacttttca | 1800 |
| | atttcttttg agatgtcttt gttgtaaaca gtattcatat gtcttttgaa agccagttaa | 1860 |
| | ctaaacagtt ttcttgagca tctttttagt ttactgaga agtattttaa attgagcttt | 1920 |
| | tctgagctcg attgcttacg tctgacacag tctcaagttt ccactgaatg gtaacaaaga | 1980 |
| 15 | ctgtagaatg ttgttggtac tgcagtgaga ggcagtcttc cttagaccag gtaagagaga | 2040 |
| | tcagtttggt tctcactgct gggtgagttt ttacagctct tattttatat tctttaagca | 2100 |
| | gcagcaatat taaattgata aatagccagg agcacgctga tttcaagacg tccttgcttg | 2160 |
| | ttgcagacag aaaaactaca gggttatgta tgggggttgg ggtggggggg gaggggaaga | 2220 |
| | attagtttat tactcagtta cttatataaa ttaattaaaa tgtgaaaata attctggagc | 2280 |
| 20 | tcagttttct taattcagga actaaagcag cagttgagga aatcagtaat tttaaaggta | 2340 |
| | cttcatgggt attacttggt aaagcaattc aaaggatagt ttttactttc atttttttcc | 2400 |
| | ccagtagtta ataaaaaag ctttgccctt aactaaacat tttttccact tacgaaaact | 2460 |
| | tttaaattgc caacagcaaa atatacttcc caaggatcct t | 2501 |
| 25 | <210> 65 | |
| | <211> 2501 | |
| | <212> DNA | |
| | <213> Homo Sapiens | |
| 30 | <400> 65 | |
| | cacaagtcaa gaccgctccc tgcttcttag cccgctgggg agccaggcca gcaggcccca | 60 |
| | cattcctgag gaagggacag ggttctggcc tggaggtct agcagaagcc accccagggg | 120 |
| | agggcccgac aggaaggaag gtaggcctgc cggaggggca tacaggagct tcctctcccg | 180 |
| 35 | ccacagtgtc cagggccaac tgctccagcc ctcaggctgg gtcaacagga tgggacagcc | 240 |
| | caggcggaag gaaacctgtg gggagggaca ccccgagac agaagcaggg acatgggggtg | 300 |

| | | |
|----|---|------|
| | gggagaggca ggaagagctg ccgggctgct gagctggcgc ctctccagca gactcaggag | 360 |
| | gggcggtgac aggaggccat tccctcctca tccccgcagc cctgggcctc tctggtcctg | 420 |
| | gccaacagta ttactatcat tattattgct gttgttcgct agcctgggcc ttagatacat | 480 |
| | tagaaaaaaa ccatcggaag atacgcatag cattggcagt ttctaaaaga attaattccc | 540 |
| 5 | ttcctgtgtt cattctgtga ttactgggat agaaatgcta tttgcattac cagcctttca | 600 |
| | ttcagttaca gagacgtgag tgctcgaagg agagacagtg atttttgcct taaattcagc | 660 |
| | ctgtccaaat cggataagat ctccgatttg ctttaagccc cgttatcact gccttcctct | 720 |
| | ccaacaacag ctgctgtgat cagcacaaa cggccaaacg ggggcaaate cgtgccaaag | 780 |
| | cagggccatg ggctttcctg atcagaaggc ctagccccag cccccaggcg cagcacacgg | 840 |
| 10 | gcggcttcct ttcagaaacc cagcctgcct cccaccagct ggagtgggtg ggtggggcgg | 900 |
| | tagtggtgcc agtttcaggg aacggccggc aaaccacact ccaggcgtgc tccagcggga | 960 |
| | gcctggagac cctaggagag ccctccccac aagcggcttc caggcaggac gcttcagag | 1020 |
| | gtcttggtcc aggggtgggg gtgaggtggg gtctaccttt gaaacagcta caatttaaac | 1080 |
| | ttcagctaca ccgagctcaa actcgattcc gcagccgagt gtcggcgcca gagaaggata | 1140 |
| 15 | aaaactcggg tctacggctc cccaccacgc ccctgggtccg gtccctctggg cttccaggag | 1200 |
| | tcctcacgcc atcctctggg ttgccagga ggaaggatgg gcggggcggg caggcgtgc | 1260 |
| | gggcgctgca gatggggagg gcgagccgc ggcacggcgt gagcggggga gaggcgcgcg | 1320 |
| | agcagggtgc ggctccgtga cagggtcccc catccgcgc cccagtgtc cccgaggctt | 1380 |
| | agtgaggcaa aaccagcaa atgtctcaga aatgcagctc agtcggtcac cgggttctgc | 1440 |
| 20 | ttcctcatca gacgcgcaag aggatggcgc ttccaatgca aatctcttgg ctccggcccc | 1500 |
| | ttggctggca gccgcgcgt ccccgctcgt cctggcgctc cgccccactcc gtggcgggct | 1560 |
| | gagacgaggc ccggcgcgga ggggacgggg cggagcgggc atccctcccc acccccaacg | 1620 |
| | tggggctggc cctccgcagt gcctgggcgc gctgcagtcg ccgcgcctcc ccggccgcgg | 1680 |
| | caccgcctct ctaggcaggg gcgggggacg aggggcaagg agtgggcgag ggggtggcga | 1740 |
| 25 | ggggcggggg gcgtcactca atcaggtggc ctctggagtt ccccggggca gggcagaggg | 1800 |
| | aacacgctgc cggggattgt gtacacgctc cactgacacc agcttcacgc tgccgggcag | 1860 |
| | tcgccgatca cgcgtggccc cgcgagccca ttggccggcg cctcacacac ctttgccgtt | 1920 |
| | gattggccgg cctcaggctc cgccccacc ccgcccgcg gcgcggggca ggctgagcgg | 1980 |
| | ctacctgaat ggggaggggg cagacggcgc tgagcgggc ggcgggcgga gcggcgctga | 2040 |
| 30 | gtgtctcgt gcgccgtct gtggccaagc agccagcagc cttagcagcca gtcagcttgc | 2100 |
| | cgccggcggc caagcagcca accatgctca acttcgggtc ctctctccag cagactgcgg | 2160 |
| | taagtcatth ggggatgccc ctgtgcttcc tcgcctggtc ttgtctgggg ggccaaaggg | 2220 |
| | ggcgcgaaac ccgagccccg gacatcagcc atgcctgaga attggggctg cagcggagtc | 2280 |
| | gtggggaagg aaaggcttc ctgcctgcag actatgggca ttagtgaggg cgtgtgtgtt | 2340 |
| 35 | ggggaggggg tcgaaccagg gggtgggat cttcagacag ggacaggggt cttgtcttag | 2400 |
| | atgtactgag gggaagggac aactccgcat ggagaccga gagggctggt gaggaggagg | 2460 |

atgacgagcg ggggaggagt ggggaggggg ccgttgccct g

2501

<210> 66

<211> 2501

5 <212> DNA

<213> Homo Sapiens

<400> 66

10 ggggctgtag aaatggcggc cccatctccc aacaacttgg gcattgtgaa tatcacctcc 60
ttaaagggga tctccttttg tcatcccgto tagagcagcc accataactt ctgagcgttt 120
attgctagct gatatatatc agaaaaatac aaattccaca aaagcagga ctggtctgct 180
tctctccctg cagggcccag gttctggcac atagttggtg cagaaagtgt gcagcctcag 240
gtcctatcca agccccaggg gcatacact cgggacttgt tctgcatatt tttacttttg 300
15 cctccactg gtactagtto ttcogtgga cagcctgagt cccttcagat acttaatgtt 360
ttttctcaag tgctgccatg aagccagatc tccaccgtct tggggcattc ctttttaggg 420
atgggaagta tatgtcgctc cttttatgtg atttacattc tatcttgat aatttgcca 480
tcacgtagt tcattcagat ctgtttgat cctgccatc tcagcttcag tccatttcac 540
tcttttaaat ctgacgaca gttacctcca acagcttcac cacaaatcac tcacaaaaat 600
20 ggccttaatc ctgaagtta ttacggaga gcacacttgc taggtgtgtg gcagatatac 660
aggaagcaca agatgaggca gcagatctag aggcaaatga cttccttctc cctgcctagt 720
ggtgactgcc agcatcacgc cctcccgga gaggtgagaa acccctccac gcaagcactg 780
gaacctcac agtcaagagt ggcaacagct ccggttactg gacttgggcc tgttgaattc 840
taatactctg tgactccaca tctgggctga atttttgctg agtatgatgg aatttacatg 900
25 cttcctccct agcccctact tgtctgtata gttggaatat ttggttcct cctctggagg 960
gatctagtag gtttagagtc tagacgctgg aactgtcaaa gttcagagga aagagctcca 1020
gctgcaaagc aagagaaatg ggctggaatt ctgacttcac cccttaatga atgcttctga 1080
tttttttttt tttttttttt ttgagacgta gtctcactct atcgcccagg ctggattgca 1140
gtggccacga tctcagctca ctgcaacctc cgcctcccag actcaagcga ttctcgtgcc 1200
30 tgagcctcct gagtagctgg gattacaggc gtgcgctacc acgcccggct aatttttgta 1260
tttttagtag agacagtttt tggccatggt ggtcaggctg gtcttgaact catgacctca 1320
agtgatctac cttcctcggc ctccgaaagt gctgggatta cagggccgag ccaccgccc 1380
cagccgcttc tgatcattaa aaaaaattt tttttttggc ggggggaacg aagtgtccct 1440
ctgttgctca ggctggagtg cagtgcagt atctcggtc actgcaatct ctgcctccca 1500
35 ggttcaagcg attttcctgc ctacgctcc tgagtagctg ggaatacggg tgccccccac 1560
cacaccagc taatttttgc atttttagta gcgatggggg ttcgccatgt tggccaaggc 1620

5 tggtctcgaa cttctggcct caggtgatct gccttccttg gcctcccaaa gtgctgggat 1680
 tacaggcgtg agccaccgtg cctggccaaa aaatttatgt tttaaaaaga ctagtcaagt 1740
 gcagtagtga gaagggggga aagagtagag caaggagtta tatctgttgc ttctgaccat 1800
 tttgaacaag ttacctaatt ctctgaggac aagctcggag aatgggagag acagttatct 1860
 atttgcaggg ttgttgggag gaataagtga catcatgagt gtgtgccagg tgtctgatta 1920
 cagaaggtgt tcaattaatc tgcaatcatt aattaaccct tcagtcgctg gtattatttg 1980
 ccatccatcc tccgagtgtt gccaaagtta ggggtgcgttc tgccagcgtc ctacgagtgg 2040
 taaggcttct ggctgccagc ggcgaaacctc tcccttcgag tatttctcct cttgctgaga 2100
 tgaaatgcga ccgggtctct ttaagggcca ggcccgaggga tccaggcggc gcccaacggc 2160
 10 tggtactagca gtcgtccgag ccgactcgca caagaaggaa ccccgggcct ctggatccgc 2220
 tcgcccggct atgctgctgt ggccgctgag gggctgggac gcccgggcgc tgcgtgctt 2280
 tgggcccggga agtcgaggga gcccggcctc aggcggcgag ccgaggaggg tgcagcgccg 2340
 ggcctggcct cccggtaacg cgctcttgg tcccgccctc caggagcccc tatgcgcccc 2400
 cctactcccg gccctcggc ttccggaacc cggccgagcc cgaagcgctt cttccgaggc 2460
 15 gcgggatttc ctccccggct gcggctggga cgggggcggc c 2501

<210> 67

<211> 2501

<212> DNA

20 <213> Homo Sapiens

<400> 67

25 atggtctcga tttcctgacc tcatgatccg cccacctcgg cctcccaaaag tgctgggatt 60
 acaggcgtga gccactgtgc ccggcctcta tcagcatttt ctttcttttt ctttttcttt 120
 tttttttttt gagacagagt ttagctcttg ttgccaggc tgaagggcaa tgggtgtgatc 180
 tcggctcact gcaacttctg cctcccaagt tcaagcgatt ctctgcctc agcctcctga 240
 atagctggga ttacaggtgc ccaccaccat gccagctaa tttttgcatt tttagtagag 300
 acagggtttc accatgttgg ccagtctggt cttgaactcc tgacctcagg tgatccgccc 360
 30 gcctccacct cccaaagtgc tgggattaca ggtgtgaaag agaccattcc cgatctcttt 420
 cagcattttc atactgaatg tccacagctg ccctgtgagg aggcctttta cccatatttt 480
 ctgactcaga gagaagcagc cacatgtccc ttggccatgg cagttaagac caactccatg 540
 gagctgggtg tcttagctca catctgtaat ccagcactt tggaaagcca aggcaggatg 600
 attgcttgag gccagaagtt caagaccagc ctgggcaaca tagccagacc ccatctctac 660
 35 aaaaatttaa aaattagcca caaaatttaa aaattaacaa caaaagggcc ggggtgcggtg 720
 gctcacgcct gtaatcccag cgctttggga ggggtggatca cgaggtcagg agttcgagac 780

5 cagcctggcc aagatggtga aatcccatct ctactaaaaa tacaaaaatt agccgggcgt 840
ggtggcgggc gcctgttgtc ccagctaccc aggaggctga ggcaggagaa tcgcttgaat 900
ccgggagtct gaggttgagc tgagccgaga tcgcagcatt gcactccagc ctgggcgaca 960
agagcgaaac tccatcttaa aaaaaaaaaa aaaaaaaagt ggaagatgag gaagttgatc 1020
10 agacatcaag gatgagcgga tgacttaata ggcttctttg ctaagacttg gctgggcagg 1080
tgaaagacaa agtcgaggag tggttatggt gtggcacaga agaagggtca gaggacggtc 1140
tttgttacct cttcatgcct gagtttcttc ctctgtgaaa tggggataat aagagccgcc 1200
atacagggaa ttgctgctag gatcaaatga gataatgtat gtgaaacgct ctggctgtag 1260
gcttctcagc aaatgggcac gacttgcgga gtggggattt gaattcacgt ctggcgggat 1320
10 gtccaagctg ctaccctgac cgctagggag cttcagagga cagggctgca ggtgatcagg 1380
aagaggactg gggcagggtg gcgaggaatg cctcccagga gtgaaggagg gggaattcta 1440
gtcagcagga tggagtcggc caggtagaaa cgagggaag gagacaggac cggatggaac 1500
ggggaagcca aagggcaggg cgtcgagggt ttgaatggtg gccggtgcag ctttgaacac 1560
cgaggtgagg acatgcagct gtgtcctagg gtcaggaccg tacacgcctg acccaattcc 1620
15 acagcacgga ggggaactcc aggatccggc cgcgttgccc acacacttcg ctctccctcc 1680
cgcctctcgc aagcccctcc cccgtctccg tccaccgagt gccagccaat agcagaagcg 1740
acagcgcacg tgggtgccga ctcagccaat cgcggtgag tgacgaatga gcccaggac 1800
caatgagagt gccgccacca tggcaaaaaa aaaaaaatcc aatggtgacg agcagggaga 1860
acagagcagc tgccaatggg cgtgtgcgtt tcaggcggcc aatgggagga ggcgtctcgg 1920
20 cgggggacaa gcagtagcta cccgcgggag cggggagggg tccgggttcg agcttgtgtt 1980
cccccggaag ggtgagtctg gacgcgggag cggaaggagc gcgccggag gtcctcagga 2040
agaagccgcg gggactggct gcgcttgaca ggctgcactt ggatgggagc acctggtgcc 2100
tcgggactgc tccgatgccc ggtgggtgca catcccagtt cccgccgttg ccggccgggt 2160
ttagaggttt tggggggagg acatgggggc gtgcagcctt cccagttgca aacttcactc 2220
25 cgaccctgtc ttcaaagctg ggtctgggtc cagtggggac gagaaaggag gaaggaggaa 2280
gtaggctccg cgaaagcccc atccccgga tctcatctat aacatgaata ggtattaatg 2340
gcaaaggcta attaaagcgt tactgtatac caggcacttt ctctgcctcc tcgcgttaaa 2400
tcctcccagc agccttttga ggtagacact gttacatgcc cttttccag atgaggaaac 2460
cagcaacatg ggtggaagtg acagcccctc cacttcata c 2501

30

<210> 68

<211> 2455

<212> DNA

<213> Homo Sapiens

35

<400> 68

| | | |
|----|--|------|
| | ggagtgaag aacacagaac taaaacagag cttgaaactt aaagaaagg agagacttgg | 60 |
| | gggaggagtg ggggtggagtg acgtgatgtg ctgctggaaa ccagcagttg gtggtttctt | 120 |
| | cttgtgcttc ctcttctgtg ggttttctcc tgccttgagg agggcctttt tctctcctcc | 180 |
| 5 | cgacagaaag gctatctttg gtgttcgttc ccttgaactg taacatcctg taagggtatg | 240 |
| | attccatgcc tctgtgtggg tgtgaattcc ctcatggtga ccctcaaaat ctgcacacag | 300 |
| | gaccccttcc cattgagggg aggggatcaa aacaactcta cttctcaggg tctctcctg | 360 |
| | ttccaactgg tctgtgtcca agagaagcct taggtaaatg gggccagctt gaagatcaaa | 420 |
| | caggtttggc agcctctccc ggctctctt ttctctccta cagctttata gctacagctg | 480 |
| 10 | ccttgatata aatattgact ttggctggct ggcatgacta cccacagggt atcgtgcctt | 540 |
| | aatttaccag gtgacaggca acgctgccct ctcttgaac catccagcag agccagggt | 600 |
| | gtaccccaa atcctgcaac agaggtttcc ctccatctca cctccctgtc cctgcatttc | 660 |
| | tcctatctca gtagctctc ttccctctc tgggcttctc ttccactcc ctcccttcc | 720 |
| | tgggttggt aaactagtcc ctaatctctt cacaccccag attggaagg gggctcctcc | 780 |
| 15 | ctgacactcc ccagagctgt caccaacctc ctccaagtt ctatagctcc attgtcaac | 840 |
| | agatttgcca ggggtaacca ttaaccagc ccttaactct gttccccac ctttcttgct | 900 |
| | ggaggggatt ttccaattac tggttagcac agctaggtea tctaccccc accatcttcc | 960 |
| | ctaacttctt gggttgggg gctggggagg aatctcccca tctcagggt ctaggaacaa | 1020 |
| | agctggggag gatggtgcat ttaaagggt tatatatata tatatatata tttttttct | 1080 |
| 20 | ttctccctca taacccacc ccgcaacac acacacacac acacacacac acacacacac | 1140 |
| | acacacacac agacgcacaa ataagcttta tggagcagtg acttcattat gttcaccgt | 1200 |
| | ttgagtccaa cccctggccc aaaataggca ctaaatagtt gccgaatgca tgaatgatag | 1260 |
| | atacctctct gtcttcaggg gtgtgtagaa gtgcgaagg gtagggcat gtcccagtag | 1320 |
| | gggtgtgagt gttctgatca gaactacttc tctctgccag aatttgatgt aattcgaatg | 1380 |
| 25 | cttccacctc tgcttgaagg gtttaaataa taaattaggc cctgtcgtgc cattatgggg | 1440 |
| | gtggtcatac cctgtaccca ggaacaggc acggtagggc tgagacagaa gtctgcttg | 1500 |
| | tttccgctta tttatttgaa acaccgctca tttaggctctt actttgtttg ccaggcactg | 1560 |
| | ttctaagctc tgtataaata ttaactcaga gggtaacaaat attaaactaa gaggttgtgc | 1620 |
| | aggaaaaaaa ataagcgct ctggctcttt aagtttgcc tccccctcaa aacccccgca | 1680 |
| 30 | acggtcccaa accccttcca gggactggga ctacggacct tggccgacc ttctcgggg | 1740 |
| | cttccactg cgccaatcaa atcccagaaa cagtgagtgc tagaggcccg gctgctaagc | 1800 |
| | aacggcagag ggcgggaagt ttgaacgttc tggacccgcc ccgaaggcaa ataggccaat | 1860 |
| | cagcgtccag actcttcagc tacggcagtc cgcttctctt cctcgccctg tcggatctct | 1920 |
| | aggctggatc cgggcctctc caatcaacag cggctaggag ggcggggcgc gtgcgcgcgc | 1980 |
| 35 | acctcgctca cgcgccggcg cgctcctttt gcaggctcgt ggcggtcgg cagcggggcg | 2040 |
| | ttctccacc tgtagcgact caggttactg aaaaggcggg aaacgctgc gatggcggca | 2100 |

| | | |
|----|---|------|
| | gctgggggag gaggaagata agcgcgtgag gctgggggtcc tggcgcgtgg ttggcagagg | 2160 |
| | cagagacata agacgtgcac gactcgcccc acagggccct cagaccctt ccttccaaag | 2220 |
| | ggtaacctcc gcgtgacagg aatgaggggtg gggcgcgtgg agtttccac aatctgtact | 2280 |
| | ttagttaaat acccgagaat tcacctctg tgtccacagc tctccacgcc cctcagccct | 2340 |
| 5 | gccccgcagc cctgtagcag aagtacttag tgctttgcat tctgcgcgcc accctacccc | 2400 |
| | ggcctcctct gtgaatcggt gcttccgaac cgcctcact ttttgcaccc gcaga | 2455 |
| | <210> 69 | |
| | <211> 2625 | |
| 10 | <212> DNA | |
| | <213> Homo Sapiens | |
| | <400> 69 | |
| 15 | ttttaaacga gaagtgatgt ttccggagca ttaaaactga agtgatttca aaaccatgtt | 60 |
| | gcactcacac gaacaggtgt gcacttaatg gactaaacta gttcagctga catgtcttct | 120 |
| | tcattaggaa cagtgtggag actgaaaaac taatttagcc tagagcagct atttaattgt | 180 |
| | aaagtctcct ttctcaaata ttgatttact atgtgaggaa atatttactt tgtatagaag | 240 |
| | tgtgtggaat tggacgaggg ggttgaccta cacatgtggt ttggtataca catatcctca | 300 |
| 20 | ttacagaggg tgtaatgaag atataggtgg ttcagcacca taggaaaggg aaaaagaaa | 360 |
| | aaaaaaagac ggtagaggtg gcctcccaag catccactcc cactcctctt gttaatgatt | 420 |
| | cacaatttgt tgttattggt gtcatttact gttctccaca cctttccaca aggccttgt | 480 |
| | gctttgaaa aatatgtctc tactccggat agaagtggg cacacagggc caggcgcgt | 540 |
| | ggctcacgcc tgtaatccca gcactttggg aggcgagggc aggcagatca caaggtcagg | 600 |
| 25 | agttcgagat cagcctggcc aatatggtga aaccccatct ctactaaaaa tacaaaaatt | 660 |
| | agcctggcgt ggtggcacgt gcctgtagtc ccagatactt gggaggctga ggcagaagaa | 720 |
| | tcacttgaac ccgggaggca gaggttcag tgagccgaga tggtagcact gcactccagc | 780 |
| | ctgggcgaga gtgcaatgag actccgtctc caaaaaaaaa aaaaaaaga aaaaaagaaa | 840 |
| | agtaagtggg gcacacgatt caggcctaag ctaaccagac caacctcatt cctgatggtt | 900 |
| 30 | gttaatgttt cagatacggg cccgcagccc tacgtagaga agaggccaag gtagaaaaca | 960 |
| | tgaatctgag gtaaaaagaa atgaggtact tgtttgctc atcaagcctc tcaattaaac | 1020 |
| | taaccttgaa gcctgtctta cctttggact tctagtgtg tcacccggta aagcccattt | 1080 |
| | gtttcaggac gtaagagttg ggttttctgt gacttggaac caaaaccatt ccaatttaca | 1140 |
| | aaatgagcaa cttaatatatt acccatgaga aatacttcat tggatatatgc tctttcctag | 1200 |
| 35 | cgtttttgaa aactaaacta ggtgggtgaa aagtatatct ttgcatgaaa ctttttcatt | 1260 |
| | ccagaaaaca ttttgtcatc ttgataataa tggccaatgc tactatatcc aaatttttgt | 1320 |

| | | |
|----|--|------|
| | cttttttttt ttttgagaca gagtctcgct ctgccgctca ggtgtgatgg cgcgatctcg | 1380 |
| | gctcactgca acctctgcct ccctggttca agcgattctc ctgcctcagc ctccctgagt | 1440 |
| | agctgggatt acaggcatgc gccaccacac ctggctaatt tttgtatttt tactgtagac | 1500 |
| | ggggtttcac ctttttgcc aggctggtct cgaactcccg acttocagtg atcctcctgc | 1560 |
| 5 | ctacctcaaa aagcaacttg ataaatccac aggctcggtta ttttttaaaa attcttttaa | 1620 |
| | atacagtata cttttctctt tttttccaga attaaccatg aatcgcacac acagccagag | 1680 |
| | gcttttaacc cgagaacgga caaaggggcc tgcttgtgca atacaattat ttttaatggt | 1740 |
| | taaacaaatt aatacataag accagcttta cctaataata taataacgaa ccaaagttaa | 1800 |
| | caacagacaa gaaaagcacc agctgtcccc gccaccccg agcgatctcc aaggggacgc | 1860 |
| 10 | gggagagcgc cgcgggggac gcggaagtct gacgtcacag gaactggggg cggggcgggg | 1920 |
| | aggccgcac accctattgc gcatgtctcc gcctccccg ccgcggcctg gcgcagtgcg | 1980 |
| | cacgcgcgcg ggtgggcggg tttgactggc cgtagagtct gcgcagttgg tgaatggcgt | 2040 |
| | tggtggcggg aaagttgagt ctctcctgcg ccgagccttc ggggcgatgt gtagtcctt | 2100 |
| | ccatagggct gagtctgga ccgaggtgag agccgccggg ttgggagtga gggagatggg | 2160 |
| 15 | aacaaggccg ccggtgggcg aggggagccg agggaacccg ggggattggg aggcttgggg | 2220 |
| | cggcgcggcc tggccgggct gggaccggcc tctcgcccta gacgcccgcg atgctggcac | 2280 |
| | cctctgccac ctctcacctg ggccccaggg gtccgcccct gggcagcctg gagtctccg | 2340 |
| | aggtagggag accgggcgga ggtggaggaa gtctttcttt ggaagacttg ctgcctgccc | 2400 |
| | agatcgatat aacatacgag gtctctcctc ccaagagtta tggcttaaaa acccctcaca | 2460 |
| 20 | aattaactac cgtaggaaat gtcaagctat gcaagaaaag ctagaaaagg ggaggggtcg | 2520 |
| | cccgttgtag catttgagc ttttctggaa caggtggtgt ttgcggagggt tgcctcacct | 2580 |
| | ccctgtagcc caggtgtctc tgcttagggc agctggccct cgcca | 2625 |
| | <210> 70 | |
| 25 | <211> 2540 | |
| | <212> DNA | |
| | <213> Homo Sapiens | |
| | <400> 70 | |
| 30 | | |
| | tagtcccagc tactcgggag gctgaggcag gagaattgct tgaaccagc aagcagaggt | 60 |
| | tgcagtgagc tgagattatg cactgcact ccagcctggg caacagaggg agactccatc | 120 |
| | tcaaaaaaaaa aaaaaatcat taaaatacag taattcaggt ttattaagtc attaccattg | 180 |
| | ggttacctca caaataaact aagtttagat gcgaactcaa agatactgag aactaatcc | 240 |
| 35 | atctcttaag ctgctaagtt agccttcttg aaacctcact tcgtagctct gcaacaatg | 300 |
| | tacttttgac atcccaagct cacaggaata aaaaaccacc tgccagttgt ttccgttttc | 360 |

| | | |
|----|--|------|
| | cacctatgtc taatttatgt acttatatth ataagaaaca aatcactaag tcttattttca | 420 |
| | tccttagtta tgttggtgtt ctatcgataa cagcatgaag atttcgggga cctggacatt | 480 |
| | aaaataagtt tgagtactgg ctttacaatc tactagggtgt gatccgaggc aagtcagtct | 540 |
| | cttcatgttt cacttctttc acttgtaaac atctattcag aagttgctgt gaacttgata | 600 |
| 5 | tttccatgct tataaaactga ttttttgaaa agagcctggt acataggacg tgataataaa | 660 |
| | tgaaagcatt tgctactttt ggaaaaacaa gcatgacaag atagtttata tactgttgat | 720 |
| | cttaagcaca gtatatgcat cttattttta gctagtctga cagtgaata ataaaaagag | 780 |
| | ttatctttga cttgcactac gagtagaaga attcaacttc agtttctaga aagatgtata | 840 |
| | agaattaaga gtggcagttc tcctagtctc aactgccatc ttcccaccag gtggttaaatt | 900 |
| 10 | cgtccagaga agaaaatgaa ttattgctat atgggattct gcagcaactt ctgtgaacat | 960 |
| | aggctcataa tttttcacca tggagactca agctttttgg agtcatagtt gtttttgggt | 1020 |
| | ctatttgtag gcatgcatcc tttgtccaga aatatacata acatttggca catggacctg | 1080 |
| | gaggtaaaag aggaggaagg cctgaggcta gacaccactc caataagtac attaatgtcc | 1140 |
| | tagaaggga atccaccttt gcagagaact cttaactatt aaaacctata gcttgtaaag | 1200 |
| 15 | cagcatthtc aaagttaaga gaagaagggt gaagggtctt gagaggctac tgactaaaca | 1260 |
| | gatgaaaatg aaggatgga gtttggtgcc aaaagaaact ccccccaaaa atcaaacat | 1320 |
| | aacaccagag taaagccctt agggcgagat aaggagtgtc aacaaaacaa gcggaaactc | 1380 |
| | gagaagcgtt aatgcttcaa agggatcaatg accacacata atctacgtag ccaacgtgtt | 1440 |
| | aaaacacacc aacgcatttt tttttcctaa acaaagtagg aaagcggact ttgcatgagg | 1500 |
| 20 | ggcgggctgc cgaccagca gtcttctctg gacagtccgt cctgattctc tctggttggc | 1560 |
| | cgtggaggga ccacatggct ccaaggctc tcagctccgg gccacacac cccgggctgc | 1620 |
| | cgcacaaact ccagccctag tctagatcca caacccttc tcgaagatca accgcgacct | 1680 |
| | gggagcccca cttcttacca tagcgaggcc ggcgatgccg cagccacatc acccttccgg | 1740 |
| | ggctcaggcg gaagaggctg catgtccctt ctgcccttct cgccctctcc agcgtccgg | 1800 |
| 25 | ttgggcttgt cacggcaccg cctaccaaga cgggcggtta agacactagg ataggctcct | 1860 |
| | ctccaccgga aaaggcggga tttagatcac gtcccgagg ccggcggaag tagctgatac | 1920 |
| | tctcattggt tgcaaaacct tgatctgtga aagcgggctt tttggaagat accggaagta | 1980 |
| | gagtcacgga gaggtaggat ccggaagtgg ggctgcctct ttaaataaca aaaatctgag | 2040 |
| | gttctgttct ttttatcttt ttgctttctt tttaaaaaag ttccctgcta cttacccta | 2100 |
| 30 | gaactccaca atgcgagaat cccctcaat ttgtgagctc ccgcgacttc ctcttggtgg | 2160 |
| | cttttgggga tgctagggtt ctgcgcatta tcctcagggt gcgacctgt caccctctt | 2220 |
| | tcagtttctc cgtttgcac tgagggatc ttgggaatgc gaagcacttt tgaaatgctc | 2280 |
| | tgtgttggtt gtgggattgg gaggacggtt gaatccagag ggtagtggt agtaggctgt | 2340 |
| | ttgagcatth cccagcact ggcctgtcct ttcaatcccc agatattggt aaactgtggg | 2400 |
| 35 | ttccaaccag gcatcgaggc tgaacgtac taggcaatth gaggtcagga aagaactth | 2460 |
| | tgtggtaacc aatgggaagg aactgccgtt tgcggaactgc agcgattgat taggtactth | 2520 |

aaagagatca actggcaaga

2540

<210> 71

<211> 2610

5 <212> DNA

<213> Homo Sapiens

<400> 71

| | | |
|----|--|------|
| 10 | ctacaggctc gtgtcaccac actgggcaat acaaaaaata caaaaaaaaa attttgtatt | 60 |
| | ttttgtagag acgaggctct gccatattgc ccaggctgga attcttacct ttgttactgt | 120 |
| | atttaacgta tctttttcct ccggccatct tcattggttt ctctctgatt tccacagttt | 180 |
| | gaatacactg catgtgtcag gcaggggctc atatttatca agttttgtgt gtgctctgag | 240 |
| | ctcaggctct tcattatttt gggaaaatta ttggaattt tctcttcaa cattttttat | 300 |
| 15 | gatttgttct ttcttcttct ttggggagtc ctattacatg catatgatat catttgatat | 360 |
| | tttcccacag ttcttggatg ctttttttaa aaaaaaactt ttttcttct ttattttcca | 420 |
| | acgtgggtaa ttctatttt tctcagctgt gttgatccta ctgctgcccc atcagaaaaa | 480 |
| | ttacctgtta tcagcgttct tcctttctta taatttgatg agtttcctcc tcatgcatat | 540 |
| | tgttcacctt tcgtacaaga gacctccaca tattaatcac agttaattta aatttccagc | 600 |
| 20 | ctgtttcaat ttctcgatca cctctgagtc tagtcctggt aattgcttag tgttattttt | 660 |
| | tgtttttgaa acagggtctt gctctgttgc ccaggctgga gtgcagcggc gcgatctcag | 720 |
| | gctgttcctt gagttcacac catccccctc aaccagcaga ttgcaaagtg tccgagtcgg | 780 |
| | gcogtgcagg agtctttgtg ggggtttcat ggactccgaa ttctcatttc tgctccatcc | 840 |
| | ccatctcatg aatccaaggc cccactctgt gcctcggtc ttcgtttggt gtgctgaacg | 900 |
| 25 | tcactacgt catctacgcc atctacgtaa tcaacacaat aaagacgcct gccgggaacg | 960 |
| | cggcccttcg gctgaatccc ttcggtggtt ccaaggccac tgccagagga tgccgacggg | 1020 |
| | tctccagggc ctctacttac ccaggacttt gaggcacatt agcttcgcct aggcactcgc | 1080 |
| | ttttacgaat tcttatgttt ggttttggtt tgagacagag tctcgctctg ccgcccaggc | 1140 |
| | tggttaaaag atagggtctc agccgggtgc ggtggctcac gcctgtaate ccagactttt | 1200 |
| 30 | gggaggccga ggcgggaggga tcacctgagg tccggagtgc gagactagcc tgggccaaca | 1260 |
| | tggcgaaacg ctgtctctac taaaaataac aaaaatcatc caggcgtggt ggcgcgcacc | 1320 |
| | tgcaatccca gctactcggg aggctgaggc aggagaatca cctgaacca ggaggcagac | 1380 |
| | gttgacgtga gccgagatcg cgccactgca ctccagcctg ggcgacagag ggagactccg | 1440 |
| | tctcaaaaaa aggaaaaaaa aaaaaagaa aagaaacaaa agtgatggg tctcgctctg | 1500 |
| 35 | ttgccaggc tagtctggaa ttctgggt caagcgaccc tccagcctcg gcctcccaa | 1560 |
| | gcgctgggaa tacaggcgcg gctaccgcgc ggtctccggc tgccgaaaca ccgccctgcg | 1620 |

cgcggaaccgt tcggccgccg ggaggaacag cggtgcccg gagctcagag gcgcgcgcgg 1680
 ctttgcgctc ccgcggcgc tctgagcctg cctcggttg gttggccagg tggctctctc 1740
 aggaccaacc ccagtcattc ccggcaggaa ccacgcttga ggggcggcag tctgcccgcg 1800
 cgagacgccc ccgcggacta caccgcggcg gcaaagcaa acgcaaaaac tacctcaccg 1860
 5 cgcgcaggcg cctccccag gaccaacatg gccacgacgc aaggcctcga cctgaggggc 1920
 gtggcctggc cgccgccagc caacgggtgt gcgcgcctgg ccgcagccaa taggaaggca 1980
 gcgcgggctc gggcgaggc agccgcggcc ggggctgtag gcgccaaggc catgtccgac 2040
 tcgtgggtcc cgaactccgc ctcgggccag gaccagggg gccgcggag ggctggggc 2100
 gagctgttgg gtaggtggc gcgcgaggcc gcgggagtgg gcgcgctccg gcccgggacg 2160
 10 gtttcgcggg ttccccgatc cttcccgcc agagcctccg ccggtcgat ccccgagcg 2220
 cgcgcccggg gggctgtgcg ggggtggcgc ccggtgggg cggcgggct gcctcggacc 2280
 cggccctcc tgcgcctggg cggacgcca ccagaccgc gcccggggg cgctccctc 2340
 tttccgaac gccgccccg ccggccgcc tgtcaggcg gcctgggggtg cgcgccctg 2400
 ggctccctc agcgagagg ccgcccctc ccagcgtcc ccgggtccc ctgcctcggg 2460
 15 ccctcctggg ccgtcttccc cggcgtccg ggtggggccg tctccgttag tttcccgaga 2520
 cctgcgcctt ggggaggagc cccggccct cttcgggagg gtgtcgttg tgggtttctc 2580
 cgcggcgtcc acctgcgcgt cgggcgggg 2610

<210> 72

20 <211> 3076

<212> DNA

<213> Homo Sapiens

<400> 72

25

gctgggatta caggcataac atggccggc cctggccatg tttttaactg tgtttctcta 60
 atagctaata atgccgagca tctttttatg tgtttcttag ccattagtag atcttttttg 120
 gtaaaatgtc tttttttttt tttttgtcc atcttaaaat tgtttttgt tttgttttga 180
 gacaggtgtc cactttgttg ccacgctgg agtgagtggt ctcaatcatg gctcactgca 240
 30 gcttcgacat ccctgagctc aggtgatcct ccacctaag tttcccgagt agatgggact 300
 acaggtgtgt gccaccatgc ccagctaatt tttgtatatt tttgttagag gtggggtttt 360
 gctatgttgc ccaggcaggc cttaacttc tgaggctcaa atgatcctcc cacctcagcc 420
 tcccaagt ctgggataac aggcataaac caccacccc agctaagatt gtttttaaaa 480
 atctttttct tgagtttttg gagtttttat gtgttaggga taccagtccc ttatgaggta 540
 35 tataattagc aagtagtttc tccactctg tgactgtgac ctttcttttt ttgaggcagg 600
 gtctcactct gttactcagg ctggagggca gtggtgtgat catggctcac tgcaacctgg 660

| | | |
|----|--|------|
| | aactcctagg ctcaagggct cctcccacct cagcctccca agtagctggg tctacaggtg | 720 |
| | tgttattgtg ccagggtaa tgttttaaat tttttgtaga gataatgtct ctacaaaaga | 780 |
| | caccatcttt gttgcctagg ctggtcttga actcctggct tcagggaatc ctccagcctc | 840 |
| | agcctcccaa agtgctggga ttacagcatg agccacatcc agcctatgat ttttcttctt | 900 |
| 5 | ttcttttctt ttcttttttt ttttttttga gatggagtct cgctgttgcg caggctggag | 960 |
| | tgcagtgggg cgatctcggc tcaactgcagg ctctggcccg cggggttcac gcctttctcc | 1020 |
| | tgcctcagcc tcccagtag ctgggactac aggcgcccgc cacatcgccc ggctaatttt | 1080 |
| | ttgtattttt agtagagacg gggtttcacc gtgttagcca tgatggtctc gatctcctga | 1140 |
| | cctcgtgatc cgcccgctc ggtctcccaa agtgctggga tcgcaggcgt gagccacggc | 1200 |
| 10 | gcccggcccc agcgtatgac ttcttaatga tgtctttgta gtacaagagt ttttaatttt | 1260 |
| | aataaagtta actttttttt aaattgtaca agcttttagt gctgtgtcta acaacttgtt | 1320 |
| | gcaaaccaca aggtcataaa gctgttctct tacgttttct tttttttttt tttttgagac | 1380 |
| | ggagtctcac tctgtcacc aggctggagt gcaatggcac gatgtcggt cactgcaacc | 1440 |
| | tccgccacc gggttcaagc gattcttccg cctcagcctc cggggtagct gggattacag | 1500 |
| 15 | gcgcacgaca ccacgcctg ctaatttttg tatttttgta gagaaggttt caccatgtta | 1560 |
| | gttaggtgc tttacgtttt ctttttagaag ttttatattt ttggctctta tatttagttt | 1620 |
| | gtgatccatt gagttgattt tatgtacgta tgtatggtcg cgttcttttc tttcctgtct | 1680 |
| | tttttttttt tttttttttg catatggata ttcaattctc ctagctccat ttaatttgaa | 1740 |
| | atgattgggc aggtactttt gagcagtga agtacagagc ggcactgcca gcagactaca | 1800 |
| 20 | cgcggtagaa agccgacctt ggtgagcgtg ttggtgctcg acagtgagca gagaaaggat | 1860 |
| | ggacgattac ggagcgccct cgtctccagt taccgcttcc tggaacacc atccgcccgg | 1920 |
| | gcggagctgt tccgcccccg tgcggtacta cgactcccag catgcacctc gcagtcggcc | 1980 |
| | ctcggtgga gcgggaaccc aggaggacct gggggtgtgg cagcaggaa gggccgagcc | 2040 |
| | acggactgtg gggccgaac tcgctcccg cccaccttcc tcgaggctgt ggcctccgcy | 2100 |
| 25 | agagccgagc gggccgcacc gccggccgtg cgactgcccc agtcagacac gaccccggt | 2160 |
| | tctagccgc ctaagcctgt ttggggttgc tgactcgttt cctccccgag tttcccgcg | 2220 |
| | gaactaactc ttcaagagga ccaaccgcag cccagagctt cgcagaccgc gccaaaccaga | 2280 |
| | ggcgaggttg agagcccgcc gggccgcggt gagagagcgt cccatctgtc ctggaaagcc | 2340 |
| | tggcggggtg gattgggacc ccgagagaag caggggagct cggcggggtg cagaagtgc | 2400 |
| 30 | cagggccctc cccgctgggg ttgggagctt gggcaggcca gtttcacct tctaagtcc | 2460 |
| | gcttctggtc tccgggcca gcctcggcca ccatgtccc ccagaccacc tctgtgggt | 2520 |
| | ccagctgcct ggacctgtg agggaaaaga atgaccggt cgttcgacag gccaaagtaa | 2580 |
| | cacggttgc ggcacctcg gtttcagcc tcaagatccc tgaaagcggg tttgcagtgg | 2640 |
| | atttaccaca acagatgggg agggactgag cttgacaaa gagccagaaa tgactggaga | 2700 |
| 35 | atgcacccct tgccactgct gcaaggggag aaaaaaggat tgatcctcag tgacaacccc | 2760 |
| | tccctcatgt ggcaggtggc tcagaactcc ggtctgactc tgaggcgaca gcagttggct | 2820 |

| | | |
|----|--|------|
| | caggatgcac tggaagggt cagagggtc ctccatagtc tgcaaggtag gcgggtcctc | 2880 |
| | cccaggatgg tcagttcccc tcttccatag ccagagaaac atccgctcct gcgtttttgg | 2940 |
| | gatcgatata attactcggg gcagggagtc ctgtttaagg cacagaggag actggagtgg | 3000 |
| | aatcatcttt gtacaggcaa atccctctct tccttacaca ctcacagagt ggcatttgaa | 3060 |
| 5 | aaatggtttc caagat | 3076 |
| | <210> 73 | |
| | <211> 2567 | |
| | <212> DNA | |
| 10 | <213> Homo Sapiens | |
| | <400> 73 | |
| | cacaccatct cttgctccgt gagtatcttt gtctctctag ctctcttct tctctcagta | 60 |
| 15 | catgtccctc cttgactccc gcctctctgc aagggtgatt tggctgcctc agttggcctc | 120 |
| | tccccctctg catctctggg tggggtgttc tctgccgtc tcccaccac acccaccccc | 180 |
| | ggtgctcccc ttccccccag caggacagcg gctcaggttc acgcacccca cggcggggcg | 240 |
| | gctggggcga cgcacgtcct tgcacacaag ccgcacgtag ctgtacttga gcacgtcgat | 300 |
| | gagcgtgtag agcgggggcg cactggccca gcggcagcgc gccagggtgca tggagctctt | 360 |
| 20 | gacgaagaag agcgcagcc gctgctggca ccacgcgtcg aagaagcggc tgaactcggc | 420 |
| | ccacgagaag aaggcccgct ccgcagctc ctgctcctcc tgccccgcag ccgtgccggg | 480 |
| | tgggggctcc ggccgctcca tctgggggc ctgcgtggag gaggggagaa caggtggata | 540 |
| | tcagacccat tcccaccg ggtatctcat ctactccatt cttggcctgc ccgctcggtt | 600 |
| | gctggtgcct ctatcgaggt gggtagcccg gggtcggacg tgccctgttt tctccaaata | 660 |
| 25 | tataaatatc aacctccatc ctatctttgg ctcctccca ccgccttacc cctggttcac | 720 |
| | ttggagcctg tcatcttgat tctaatcc aactcgtctc ctctccgca gatgtgacct | 780 |
| | ttaggtacag ttggaatctc tcttccaaa atacgacct taagctcaga tgttccttaa | 840 |
| | ggacatctcc tcaaattgtt tctcaaattc cagctaaaac ctctccccc tccagctgtg | 900 |
| | tctctcacc aagagtaact tctaaactc gtattcatct ggaactcctc ctccatgtg | 960 |
| 30 | ccaacagttg gctgtaacct ctccaaagac gctccatctc cagatgtgct ccacatcca | 1020 |
| | ggccacggac ccctcaccg gtcacatgct tcatgcacct gtggctccgc actccccaga | 1080 |
| | tgtgcctctg gcgtgcagct gttgcccctt ccccgatta tgacctatg gctcgccaca | 1140 |
| | tgcagctgta gctggggctt cctgagaca ctctcatctc cagatgtact cccacatgc | 1200 |
| | agttatccac gcttcgcta cagggtgtgtg cccacttgt ggctagtctt cctcggaagt | 1260 |
| 35 | gtcaccagta ttcacctgtg gtcccctcct cctcagatgc ggccccagt ccagctgtgg | 1320 |
| | gcccctctc ccagttacat ccaccatccc ccgcaatatg catcttcgtt ctagacatgg | 1380 |

cccctcgtcc tcggatgggc tccttcaccc cagatgctcc cccacgtcc agctgcggt 1440
 ctcccctcga gcagcccat ccagcccgt cccgacgtc ctactcccc cctccccgcc 1500
 cgctgcggca ccttcagcc ccgccgtccc acctagctgt gcctctcccc tcccgaagat 1560
 gtgcaccctt ccgcccctc cccactcacc taccgcccc ggagcggcgt ccacctccca 1620
 5 caatgccccg cggccaggcc tggcccgcc cttgctcccg ggatgcccc cgcggtctcc 1680
 cgctctctt cccgctgtc ctgcggggg cgcttcacc gattctctct ctttcctgc 1740
 cagtcactcc tcagaccctc agccacacc gctcatccag ggcgaggaa agcgcgggca 1800
 ttttccagt gtgtctgcg ggagggtcg cccacttca ccccttttc cgccctctc 1860
 ccattcgga gactacgact ccagtgctc tccgcgcac ggcgcggtg cggacggtg 1920
 10 ccagggtccc cccctaggct ctgccccgcc ccgcccgcga gacgtctgc cgcgatgcc 1980
 gtggcgcgaa cttgggactg cagaggcgcg cctggcgat ctgagtgtg tggccggca 2040
 gcggcgcgcg ggaccaacgc aaggcaagtg gggccgtcc caagcagatg ggaggcgag 2100
 ggcgcggggt gcgccaatg cttggggcct atgcttcgcc atgctgggt gtctgcagag 2160
 gagtggcggt ggggacgtg aggtgcga gagcggtg gagacggaag agcggggt 2220
 15 gcgggcccgc ggagagtga gagaggtgc tcccagagg aggggggcca ggtagagggt 2280
 agacgagaga cagagacagt tggacaggtc ctctgagaag aggccttgag gtgcgagttc 2340
 acctggaagg gggagaggcc aaatggaact gagggcggg gcggggggg ggaactgt 2400
 gtggcgggg ccagctgga atcggaagg cccccaggg gcgggggcta tctggaggg 2460
 ggaggggctg aaggagcta agggcgggg ccggggaaaa gattgcgtgt gggcggggc 2520
 20 acctggaagg gggaggtgcc aagggtggg ctggctggga accggaa 2567

<210> 74

<211> 2278

<212> DNA

25 <213> Homo Sapiens

<400> 74

tcacagaagt caaagctcag gaaaagcccc tcgagggttt ttgtgcggca gaggtgggt 60
 30 gtgggtggg attgtgctg ccacagtga ggggccctgc agaccagat aaacctcaa 120
 gtggccagaa gcgggggatg gctctgctg gtgctggggc tgccatggg cgtgggagcc 180
 agcagtgtgc ccagctccct caggcccgt cccctaggcc cttccgtcca ctgggccaag 240
 caccgtccct gcccctccct aggggcatg atctgacttg agaggttgtg agagcttaca 300
 ggcgtgggc cgtcggggag gcctcagaag cgtaggacgg ctgcgactg ccgggccgtg 360
 35 ttcagccctg gtctggcctc ggcctctaga ggaggctgcc tgcgctccag caggcccaac 420
 ccagaacgtg ggcgagctcc cttcagcatc cctgggcgga aagagggatg ggggctctgc 480

540
540
600
660
720
5 780
840
900
960
1020
10 1080
1140
1200
1260
1320
15 1380
1440
1500
1560
1620
20 1680
1740
1800
1860
1920
25 1980
2040
2100
2160
2220
30 2278

tgcagaggca gaatccgcgc cgtccctccc ttccttcccc cgaccagcct gtgacaacccc
cgggccagggg cgggggcctc cgcacaagcc tggcgctccac ttcctggata aggactcccc
ggcccactcc ggaccagggc tggggcggcc tcccaggcgc tcaactccgct ggcaccccac
cggaaaacac gtctgcggcc cgccccctcc cccaaagcac gaccactccg cccgggcccc
tcgaggatcc actcaggttc acgacgggcc cgtcctctcg gtggtctgac caccggctgg
tggagtgggc tctggggccg ccaggcgacc agggcgagcagg cgggggaggga cagctcattg
ggagggggcg cggggcacag tggggggctc gccccacccc cagggtgcccc ttccccgctc
tcgcctcgca ggcaccgcat cggggccggg aatcggtccg gacctggcgg tgggcgctgg
gaagaggatc cacctccacg tggcccggcc cgccccgggg gcgcagccag ttcccggcgc
tcaactgcccc ccttctcccc gcttccgtcc ccttctgcgc aggcgcgct ccgccccggt
cctaggggtg cttccgtggt cggcggtgc tgggctccgc gccggggtcc gagtccacg
aagccccggc ccgagccgccc ggtgcccgc gcgcagcggg gccaggtga gcgcgcgcct
cgggccgccc gcggaacaga cgcgccacc cccaggcgca gcagcgagcg cgcccgcggg
agcgggagtg ccggggacgg gcgtagcgcc caccgcccc aggggttcggg gcagagccag
agcataggcc aagggccaag ctcgggccga gagcagtggc cgcagcgccc gggggctgaa
cccacggcgc gctggcagcg cgggcccagc tgcggagacg gtcacgtcag cgtccgttcc
aggccgactg gcagtctccg ttctacatta acgtcagcac tcccgttaaa aataatgcat
ctctcccattg ccaggaggac ttaggtgctg ctaaagacca gccctccggg tgctgccagg
ccggcgctca cccgccacct tcattctccc ttctcctttg ccccaggaca gccgaggatg
tgtggttagg ttccccctac ccatggggag gccagaggtg ggaggctggc ggctgctcg
gtctcagcag accctcctag tccctcagga gaccttgctt ttgccccact tgctcgttat
ccagcctggg ccatgaagca gaggacagtt agggaccctg agcacgcggt ggtcaccccg
gtgctcacc ctcctgtgt gtccgacctt ggccctgcta agatcctgtg ttttgaattc
tggcaagggt tggatgaaag ggcagggctc cagaaaccag ctacagcgtt tgcttgggac
ctgcatgatg agtgggaatc ggagggcacc agccctgctg tcccaggctc agggcccat
ctgctcccca ggtcatgcag cctgggcccc catgccgtgc agctcgaca tatgtggggc
agagcagcca ccctgcccc agcagcagcc gtccatcgtc agacgtgatc atttctgag
gcctcgagtg tgtcagggtg tttgtgcctc ataacaaccc acaggatggt ccccccgct
ttgcagatga agaaacaaa gcagggtggt agatccagtc cttgcacttc ctgagcctga
ccttaccaca cagctgtctc ctattcggt gcttatttat ttttttccc attacagt

<210> 75

<211> 2401

<212> DNA

35 <213> Homo Sapiens

<400> 75

| | | |
|----|---|------|
| | tcattgctgt aatcctaaca ctttgggaag ccaaggtggg aggactgctt gaggccagga | 60 |
| | gttcaataact agcctgggca acacagcaag atctcatctc taccaagaaa aacaaggat | 120 |
| 5 | agaggagtca actgaaaaag atcccagtga ctaaagctcg aacaatttta gcaataaaat | 180 |
| | aaatacgcac gatataaata catggctgaa taaataaact ggggagaata gaaaaatata | 240 |
| | ctgtgcagaa gaattccaag taacttatat agatatttta cctttacctt caaggaagta | 300 |
| | gaacataact tttcattcct tcccaggatg ggctaggcat gatgacttcc ttccaaagag | 360 |
| | tacagaacgg aaacagggca gggggattaa cagtggagaa acctgaccaa cgctactgca | 420 |
| 10 | gctagggtgat caaggccaaa acatcgacag tgataaagca tgctgagagc acctttgatt | 480 |
| | tgatgtagt aaaatcgtgc tttacctctg taatcttctt gccaaaaacc cataatccca | 540 |
| | gccccatta tgagagaaac attaggcaaa tatcaattga gaaatattct acaaaatacc | 600 |
| | tgactggtac tcctgaaaac tgtcaaggtc accaaaaaca ataaaagctc aagaaactgt | 660 |
| | cacagcccag aggaacctaa gatgtgacta ctaaatggca tgtagtacct taaatgggat | 720 |
| 15 | cctggaacac aaaaagagta tcaggtaaaa actaagagaa tcagaataaa gaaaggactt | 780 |
| | ttgttaataa tagtgtatca atattgggtc atcaattttg acaagtgtac cataactaata | 840 |
| | atgcaagggtg ttaataagaa acatttcagca tgagattttt aggaattttc tatattatct | 900 |
| | tcacaatttc ctgttaatct aaatctctcc taatgacaag tttattttaa aagtaaaaca | 960 |
| | aaacttgaag gagggaggaa acaagaaggg aggaacatt ggagacagaa ccagcttggc | 1020 |
| 20 | aagttgacag ataaggctctg agaagtaggc aggggaaaga tcattcattt caggcaatat | 1080 |
| | ttttccattt tacctgtata agaaccatat gagccctatt tttctttctt tctttttctt | 1140 |
| | ttctttcttt tctttttttt ttttttttgt agagatgaag atttcaactat gttgaacagg | 1200 |
| | ctggtctcaa actcctggcc tcaagcaatc ctcccacctc agcctccaa agcatgagcc | 1260 |
| | accatggtgg gctgtatga aggaactttt taaaaaatgc tacaagccgg gtgcagtggc | 1320 |
| 25 | tcattacctg taatcccagc attctgggag gccaaagtaa gaggatcact tgggccaga | 1380 |
| | agttcaagac catcctgaac aacatagcaa gacctgttc tctgcttaa aaaaacaaaa | 1440 |
| | acaagctggg cgtggtggat cacgcctgta atcccagcac tttgggaggc tgaggtgggc | 1500 |
| | agatcatgag gtcaggagtt cgagaccaga ctgaccaaca tggtgaaacc ccatctctac | 1560 |
| | taaaaatata aaaattagct gggcacggtg gtgtgcgcct gtgatcccag ctactcagga | 1620 |
| 30 | ggctgaggca ggagaatcgc ttgaaccggg gagacggagg ttgcagtgag ctgagaaagc | 1680 |
| | agtgcagtga gatagcacca ctgtgctcta gcctgggaga cggagtgaga ctctgtttca | 1740 |
| | aaaaaatcag cctgccagct cagagcgctt cagcgccgtg ctcgggacat cccgccctgc | 1800 |
| | ggccagcccc cgcgtgacgt caccgcattc cggctccgct cctcccgccg cggcgccgc | 1860 |
| | accgcagtga cagccagccg ggcccggtgc cggagaggaa gtgcggtccg cgccaagccc | 1920 |
| 35 | gtcccccg acgcccgtc cccgcggtc ggggtgacag gtcgcggccg ccggacgcag | 1980 |
| | cgcggggcag gcgcgggcag agccgagcgc agcggaggct ccggcgaggg cgcggggaaa | 2040 |

atggctgatg acttttgctt cttctcgtcg tcggagagcg gtgccccgga ggcggcggag 2100
 gaggacccgg cggccgcctt cctggcccag caggagagcg agattgcagg catagagaac 2160
 gacgagggct tcggggcacc tgccggcagc catgcggccc ccgcgcagcc gggccccacg 2220
 agtgggggtg agtcagcgcg gggcctggag aggggctcag ggcgcgcacc cgggggaccc 2280
 5 cggccggggc ccaggggcac aggaagaga gcctgctcta ggccaccggg ggcaggagct 2340
 gggagacgtg ggaagaatc ttcttgaga tctccatgta ggacttccga gctggggatg 2400
 a 2401

<210> 76

10 <211> 2501

<212> DNA

<213> Homo Sapiens

<400> 76

15

ccagcctggg ccgcagagtg agaccctgtc tcaaaaaaag aacctactag tctacatacc 60
 acacttcctc atccccatct gagactatat atattttttc taacatgagg caatgccaaa 120
 aagaggggct ggtgagtga agtaagaaca gaaagacatg gaggcaagtc ttatagaata 180
 atagccaaca cttaaaacta cacttaacag cgtgataggt attgttccaa acacattaaa 240
 20 ttcatttaat ggtccttaca tgtctatgta tttggtgatt attatcctta ttattcacat 300
 tgctgagtgt attattctgt tctcatgatg ctgatagaga cataccgag actggataac 360
 ttattaaaaa aaaaaagggt taatggactc acagtccac gtggatggg agtcctcaca 420
 atcatggtag aaagcaaaag acacgtctta catggcagca gggaagagag agaaatgaga 480
 accaaacaaa aggggtttcc cttataaaa ccatcagctc tcatgcgact tattcactac 540
 25 catgagaaca gtatggggga aaccaccccc atgattcaat gatctaccag gtgcctccca 600
 caacctgtgg gaattatggg agctacaatt ccagatgaga tttgggtggg gacacagcca 660
 aaccacatca ctgaggaaac tgagtatatg ggagattagt aacgccaac acagctggta 720
 ggtggtggag ccaggcagtc tgactctagg gtctggactc tgaactgcat catgctgcca 780
 agaagttcct ctttttttcc tctctctaag tttcccttat tccctacag tcattccttc 840
 30 aacagcattt ccttcacat cttttctact tctactatat aattaatttt ttcttcttgg 900
 tcccaaattc caacgtgcaa atgcagcctt atatacccta attcatcttt accttagac 960
 tttcttccaa tgtttctact tcattccatt ttaaatttat ccatgagatg cctatttaca 1020
 agctgtaacc atcatgaagt gaatgaagaa taatacctac tactgtacaa tagaattcca 1080
 agagtataaa taggagttat ggctttctga cttgaaacta aatacttgat acttgatttt 1140
 35 gctgtctgag atcaatctga aaagtaataa taactactaa catttggtga gcatcaattg 1200
 tgggccaagt gtcatttcaa tcactctgta catattaact catttcatcc tacaacaacc 1260

| | | |
|----|---|------|
| | cggtgaggca agttctgtta ttctgtttta cagttgagga aacagaggca tagagagctt | 1320 |
| | aagtagtttg cccagtagat agccagaaga ggagccagga tgggtctcgg gcagtttaac | 1380 |
| | agcacagctg aagtcttaac cactatgcc aacagctttt ggtcctacac atcccatggg | 1440 |
| | aagaggaaaa taaaaggtta tctatttgta taccttttta tttctgatat aagaagcaga | 1500 |
| 5 | attcctttca catgacctat gtctatttaa tacgtcattt tgaaacttac caataaaatt | 1560 |
| | tccaagcgc cagaaaactg ttagtggtt tttccatttc tctctatttt tttttgtgct | 1620 |
| | actaattttg cttctttccc tcagaaggct gccggaatag taaacattca ctgacatgtc | 1680 |
| | ataattactg gaaaatgggc actggaaaat cacattgtaa ttaattcaaa gcatgttttc | 1740 |
| | caaatgtact actttaaatt ggagcttata tcataatcca aggaaacctt tgtgtgtgta | 1800 |
| 10 | ctgttccac attgctcagc ctgggatatc caggagtaat tcaccttgcg cctgcctcca | 1860 |
| | gaccatcttc catggaaggg ggtgacctt tgctcttgg caaccactat ttctaagctg | 1920 |
| | ccaacattac tcttgcatca tcaacattct aacttcatgg gaagggtgt ggtgagtttc | 1980 |
| | tggaatgtga ataggaagtt gtttttctaa acagcctgac actgagggga ggcagtgaga | 2040 |
| | ctgtaagcag tctgggttgg gcagaaggca gaaaaccagc agagtcacag aggagatggt | 2100 |
| 15 | gagtttattt ttttctgcat gggaagtgtt tgaagtgagt tggagtggta tggagtaaag | 2160 |
| | tcaggcaggt aaaggttcag aaagtgagga acagcgatag ccatggagtt ttatgttgaa | 2220 |
| | ttgcctatta gattttgtga gtacttttaa acttgctgtc cactttgacc ctccaacac | 2280 |
| | ccttgtagt tgaggttgc atttctattt tacaataaaa gccatcgtgg ttacagagg | 2340 |
| | ctgtgtttta tctaagcttc actgttaggc tacatgatgt tgggatctgg ggctgtcct | 2400 |
| 20 | ctggctcgc agctgctgtt cctcctacta gaatttatag gggctctctg agaatagatc | 2460 |
| | atggtaaacc tgtcacccca ttttccaaga ctgtacttct c | 2501 |

<210> 77

<211> 2501

25 <212> DNA

<213> Homo Sapiens

<400> 77

| | | |
|----|---|-----|
| 30 | cctgggtcct ctcttcacgc tcccaaatg tactctattt ttatctgttt cacgaacgct | 60 |
| | ggtccagata gtcttccatc cccactgac tgttagaagt gactctcagc ttttgtccat | 120 |
| | ctcgaagttt ctgtgctcag tgtgcctctc agactaaagg ctctcttgg gaagccccga | 180 |
| | ctctcgcttc tcaggacaga gatccagggg ttgggggagg aaaaggttga ccagaagcca | 240 |
| | tagcggagca gggagagaga gtgtgaaaga cagaccgcg gccaggctcc cagttctcca | 300 |
| 35 | gctcgtagag ggccaagtg gccgtataa tctgaaagag cagatatcgt aatcccatag | 360 |
| | tacttcctat tggctgcagg acacagttct gtcctgacac tgaaatttgg gtgtgtcagg | 420 |

| | | |
|----|--|------|
| | gttctgggaa ttcacaacgc tcacaacttg tgaagcagct gtgggggtggg ggatggggag | 480 |
| | ggtttcagca gaggaagtga ggtcagtc aaattgatgc ctgtctgagc ttttagccat | 540 |
| | tatctcccc agcctctatt cctgtcaaaa ggtggggcgg ggcaggagga ggggtccctg | 600 |
| | gctcatcttg tagaatcccc atattagagt aagacacctt agaggtctac tcctgcttct | 660 |
| 5 | aatacccacg tctttccaag tgtctctgag gccacccctt cccagcctt ttcatttatt | 720 |
| | catttaatta acgaacgcct tcattgaggg cctcctctga gtcaggctca gccagccagc | 780 |
| | atctttgcta tgagctgaga taagcatcat ttccgtctat tctcacaacc accctatgag | 840 |
| | gctggcacgg tttactatgc ctatttagca gatgggggac tgaagcatgg agaggtgtca | 900 |
| | ctagcctacg gtaacacaac cagcctgcat tcctagtagg tagtttgact tcagagtctc | 960 |
| 10 | tgtggataac caggaggcta ggactaagac cagagtcctg cagggtactta gatggttga | 1020 |
| | gcaaagcagg gcagtgaggt cagtgtctcc agcctgtgca ggagcatcag gaagagtctg | 1080 |
| | tgtccccctc ccctgccggt atgaagccat tctgcttccc tccccagctg ccttggtgca | 1140 |
| | gcagagttcc agggaggctc cattccccac ctctatctaa agctccattt gctgggggtg | 1200 |
| | gggccctgcc tggaagggga aggtccaagg ctgctcccag cgtgtccctc catcctgact | 1260 |
| 15 | gtccctggcg gggcggggggt gtctttgtca cccagctgca caacggccag gaagggctca | 1320 |
| | aaccatcctc agggctaacc caaggccgtc ctctgggctt gtataccctt gtctgagtg | 1380 |
| | cggatcgga gaggtgctg aagacaggag gggacaaatg ggggacgaag gggcccagg | 1440 |
| | gaggggactg aaggatttg gccaaagtcg gagttcccga gggcgagtc aaaacgcac | 1500 |
| | tggattttgc tagcccaaaa ctctgccctc attgtgcaa gcctcctaga ccgaggacct | 1560 |
| 20 | ccgggctgag ggtggggtaa ggataggtag tgtccctccc cgtccacccc ccgcctgtcc | 1620 |
| | cttctcgggt ggcccttcc cggcgccccc attccaggcg gccctccgc tgctgccagc | 1680 |
| | cgatccccct ctacccccac ccaactactc ggccgccaga cgttgctac agtctcggt | 1740 |
| | ctgtctcca cggtgtggg tccggacccc acgggacccc tatgggaccc ccacaggacc | 1800 |
| | cccacggcct gagtccaagg cccgccccct cggggaggcg gatgtgggag gcccgggccg | 1860 |
| 25 | ggtgcgggcc agcgacccgg gagctgctgg cggtgggag gggaggccgc cctgaggggc | 1920 |
| | tgggagcggc gcgggggtgg gtcccgtcc tgcagcccca gcgagggcg agcgcgggcc | 1980 |
| | agtcggcgag ctgggcaata aggaaacggt ttattaggag ggagtgggtg agctgggcca | 2040 |
| | ggcaggaaga cgtggaata agaaacattt ttgtccagc ccccatcca gtcccgggag | 2100 |
| | gctgccgcgc cagctgcgcc gagcgagccc ctccccggt ccagcccggt cggggccgc | 2160 |
| 30 | gcccggaacc cagcccgccg tccagcgtg gcggtgcaac tgcggccgc cggtggagg | 2220 |
| | gaggtggccc cgttcgccg aaggctagc ccccgccacc cgcagagcg gccagaggt | 2280 |
| | gagtcgaggt ccgcgagcg gaccgggtg cggcggcct gacccccgt tcagtgggcc | 2340 |
| | cttcttcgg gcggaccca gagtcaccgc agagtgtcg cgggaggctc agtcccagct | 2400 |
| | cattagaaag gcaagctgct cctggctgac cagcacagc tcccatgacc ctacctgaga | 2460 |
| 35 | cttgaggggg aatggacgag actggactgg aaatcagaaa c | 2501 |

<210> 78

<211> 2501

<212> DNA

<213> Homo Sapiens

5

<400> 78

| | | |
|----|--|------|
| | tggetaatttt tttgtatttt tagtagagac ggggtttctc catgttgagg ctagtctcga | 60 |
| | actcctgacc tcaggtgac tgccgcctc agcctccaa agtgctggga ttacaggcgt | 120 |
| 10 | gagccaccac gcctggcgc taactacatg tggtctatga ggtgaggtcc ttcccagacc | 180 |
| | ctggaatcag ggggtgcaat tagggccaa ataatgaggt tggactacag ataaccatc | 240 |
| | tcctttctta cctttgacta gatccaagga ctaactcca agaaccgag catctgtccc | 300 |
| | caaaactgaa aggattggac tagtcacccc ttgtttccct acagccacat cccaggcacc | 360 |
| | tggcccttgc tttgtccaga aattcagcta taactccaca catctgatgg ccctttctgg | 420 |
| 15 | caagcaggca tttccatcag gaccctcagc tgccagacac atttactgga ggtcacttat | 480 |
| | taaacctggg ctcaatttcc acacagggag gctactgaag catcacactg ggtctcccag | 540 |
| | ccccttctca tagaggaaag atctctctgt cctgcagggt tggcagtcag cgccaagtaa | 600 |
| | agggaaattta gctcttggcc caagatccct gcccaggaaa ggtacttgcg cctgctggaa | 660 |
| | actttgggct gaagtatact cctttccaaa aactcaggtc tgatatttac acaaagtctg | 720 |
| 20 | aaattaatgc agagaaaact tccaagtgct tggactggag cagaaggctg agaacaggaa | 780 |
| | ggggctggtc cctggtacta gttttggttt tttgggtggt ttttttttctc ttgttttttc | 840 |
| | tcacagaaca gggcaaagct gagtgtccct ggatgagtga agcaggagga ttaatcatgc | 900 |
| | ccagtgtctc tccactttaa actggttttc ctgggaattt gcaattgaga gtggggaggg | 960 |
| | gtaagaatcg tgggaaaagg ctgatggtgt tcagccaaat tcctccttca cgtgcccacc | 1020 |
| 25 | cttctacagg cacatgcttt ggggccatcc acggctgcag ccaccccatc cttaggaagc | 1080 |
| | accactggcc ttcctttccg gtacctggac tcagcatcac tcccagcctc ttggagatgc | 1140 |
| | agccttcatt cagcacacag ctcagetctg agttctgttt ttgtccctag atgtctctgg | 1200 |
| | ggtcacctac tactccctgc ttggtggccc aggcccatcc ttctccactc ttgcacctct | 1260 |
| | tttagcagaa aaggagttag aatggatatt tccatgggcc gtgtgtgcac tcccggctac | 1320 |
| 30 | ccctgacagc tctactcaga gctaccctcc ctctggggc ttcttatgtg ttctaaggct | 1380 |
| | gaggcaggaa gactgtgaga tcaggtgaca ctcaacagtt atgatcggtc ttaagattaa | 1440 |
| | cagtctggc cgggcgcagt ggctcacgcc tgtaatccca acactttggg aggccgaggc | 1500 |
| | aggcagacca cgagatcagg agatcaagac catcctggct aacacagtga aacccgtct | 1560 |
| | ctactaaaaa tacaaaaaat tagccaggcg tgggtggcgg cacctgtagt cccagctact | 1620 |
| 35 | caggaggctg aggcaggaga atggcgtgaa cccaggaggc ggagcttgca gtaagccaag | 1680 |
| | attgcgccac tgcactcccg ggtgacagag cgagactccg tctcaaaaaa aaaaacaaca | 1740 |

| | | |
|----|---|------|
| | acaacaacaa aaagattaac actccttcta ctccaaacc taatacaaag ggacattgcc | 1800 |
| | tagtgattaa gagaattcat tcattcaaca aatacttgtt gagcacctac tatgtgccaa | 1860 |
| | gcactgttct aggcaccgga aatacagcag tgagaaaaac caaaaaaact ccctgccctc | 1920 |
| | atggggtgta tattcaagta gctgaaacag acagtgaaca aacaaaaaag gacaataatt | 1980 |
| 5 | tcaaataata atgatgctat cggccagggtg tgggtggctca tgcctataat ccagcattt | 2040 |
| | tgggaagcca agtcaagcgg attacctgag gtcaggagtt caagaacagc ctggccagca | 2100 |
| | tggtgaaacc ccactctctac taaaaatata aaaattagcc agacatgggtg gcacacacct | 2160 |
| | gtaatcccag ctacttggga ggctgacgca ggagaattgc ttgagcccgg gaggtggagg | 2220 |
| | ttgcagttag ccaagatctg acaggccttc agcaccactg cactctagac tggctgacag | 2280 |
| 10 | agcgagactc tgtcaaaaaa aaaaaagcta taaatagact ttaacagggt aacatgatag | 2340 |
| | ggagggaggg ataggggagc aggggtgtca aggaaggagc atttaaacag gctagaatga | 2400 |
| | caatggccag cgagggaaag atccagaagt gtgtgctgga agaagaaaga gcaagcacia | 2460 |
| | aacccttagg acaaaatcag ctogtgtgggt caaggcacag c | 2501 |
| 15 | <210> 79 | |
| | <211> 2501 | |
| | <212> DNA | |
| | <213> Homo Sapiens | |
| 20 | <400> 79 | |
| | tgtttctgac ccctggctgc agcctaattgg gccgactgct ggacagcgggt cctgagtcct | 60 |
| | gtttgaattg gtgctgcccc gacatcctct gacctcagct aatgatcctg cctgccgagg | 120 |
| | gcagacagggt ctctgcaacc ctatgggtgg taggggtgggt gatgagagga gaggtagtct | 180 |
| 25 | cacttgacaca gatttttggtg tatggttctg tcttttgac tctttcaaca gaggtctgtc | 240 |
| | cagtccctct tgcaagtgtg gggaggggggt ggtgcaggac tatgaggtaa ctgtgagaag | 300 |
| | aggggctcca gcagaaccag ggtccaatgg ccttgaagag atggctgggg acagctggac | 360 |
| | tcattacgtc tactcctaaa tggaggaaac gaccctcag ctacacagca cctgagccag | 420 |
| | aatgtcacca tgggtgctgct ccacaggatg acagctacct ggtttgtgag ggccccatt | 480 |
| 30 | ctagggacag ctacttcatt ctgccctccc agagcagcaa gcaacaaccc tatgccagga | 540 |
| | ggccaattgg cacgtcaagt gccagctcca atcgattgat agtagctgcc tggctctgaa | 600 |
| | aggcagctgg gatcgattca ccatgctgcc agcacacaga tggaccagc ggtgggtcca | 660 |
| | gcagttagtt cttgccttgg gccatttcat tttctttgtc ctggccaagg aatgattgga | 720 |
| | tgaacacact ggactcccaa tatgggtgga taagacaaga gtgtctgtgc acaccctcc | 780 |
| 35 | accactcata agcatggttg tgggcagttt ggttcccag gcggccttgg agaattgcaat | 840 |
| | gagccgagga actggctcatc tccagggtgca tccagggcag gaaaggatga cagcatgcgt | 900 |

gagccagggt cactggctaa gaagtcattct caggacctcc cctagaaaa gccactggg 960
 cagcatccct gctgggtccc ccctacacca caagggtacg cagagctggc ggaggggtcat 1020
 ggtcccactc atgtcagggt ctcttaattct ggcaaggaaa tgtaacctac gtgaattctca 1080
 acaggcagtg aagcacggtt tcttctgac tccaggtagg gtgaagaaaa tgggacagta 1140
 5 gtacgggggtg cgggcataaa cgcacaactc tgcctcccca gacgcagagc tgtggggctg 1200
 tgagaatgcc aggaggaggt aagaaagggc ggcccatggt ggggcctgca ggggtgggaca 1260
 agcccaagag gtctctacat ccaggcctgg tgggggaggt gagccctgg tttaccgagg 1320
 gggctccctc ctgcccctgg aaatactgca gtcctacct ccactgtctc cccgctgagg 1380
 ggacccaagg gcgtgaggat gagagagccc ccaggcccca gggtcagacg actgtgttca 1440
 10 agcaagttag aacctctctg aggtgtttc ccaactgtaa aatggggata gcagcagaac 1500
 tctctctcgc ggcttgctg aagaatacaa ttgatgtcg acaggagga gcggcgcgca 1560
 gcgcgcagcg agtagcaggc gctgaagaag gatacctgtg aactgggagt ggtggcgag 1620
 gctacgcggc cagagtccgg ggaaggggag ccggtctgac cagtcctgac tcggggctgg 1680
 atggctgggg gatgttctcg taagtccggt gggaggagag ggtcccgcgt accctgccac 1740
 15 cgccgcgcga gaggttcggg cagggtcggg gccgcggccc ctccgcgagg gggccggtca 1800
 tccgcgggga ctgacatccc ggaggcccaa tggcaagcgt tcatctccgc gcatccgccc 1860
 aatcggcgcc ggttgccgtg ccgcgcgggg tctctcgacc aatgggaaa tttgctgtca 1920
 gatggggcgg ggcggagatt cgcgtcgccg gcccggtccg ctttgcgac gggccgcgtg 1980
 agggcgggag ggctggcccg ggggtctcgg ttgcgcgctg ggcctggagg gagggggcgg 2040
 20 ccccgccacc ggtccgagtt gggcccggt ggactgcgac ccgcgcgcg ccgcaccgcg 2100
 ccgcgcctg ggaacgcgc tcccgcgcg ccaacggacc cggggaagcc cttctggggt 2160
 ccgagggcgc gctgcggggc cggccacgct gcgtccagg taagcctgag ccagtgggcg 2220
 ggggtgtgga cccggggctg gggcctcggg tcggagccgg gactgggggc ggggctgcag 2280
 atatgggacg cattccgggc agcgggtccg acagggtcct atccctggag tcgagatccg 2340
 25 ggcgaggggtc tgggcccggac gtcggagcca atctccgcc caccgcgctc ttgtccgcgc 2400
 gctctgcggc gtccgagacc ccgggcccgc gggggcggt ctctttgtgc gtggccttgg 2460
 ggccctaccc taccggtccg ggcgtcttgc actgagcact c 2501

<210> 80

30 <211> 2501

<212> DNA

<213> Homo Sapiens

<400> 80

35

acagatgacc gaggggctcc cagcccgga ggtggaaatc cagcaggat ttccaaggcc 60

| | | |
|----|--|------|
| | tagtttgcag ggctccagga tcgttcctag atcctgggtct tgcagccttg acaaggggaa | 120 |
| | ggagggaggc agcagaagga gggcagaaca atccatgccca ggctgtgatt tgccaagtga | 180 |
| | ccatctggga agaattgggt ctcagaccag ggacaggag cagaggcaag cccgcatctg | 240 |
| | ccctgggtgc agaaccggga ttccagactca gggccccgat ttctgcctgg atcgctccac | 300 |
| 5 | tgggaggagg agtgactgtg gacacatcca gggttctctc caagtcggct tcctcatctg | 360 |
| | ccaaatagag accgcagacc accagctccc aggcagggtgc tactcttcg gcccctocca | 420 |
| | aggcaggagg gccaggcgta ctcgagacac aggtgtgctg ggggccagg tgggccagcc | 480 |
| | agcagcatcc tgcagggtaa tgggagcagg tgggcacccc gaggtggca gtaaactctg | 540 |
| | gctatctgcc cccaggctcc caggaggggt cttgggcctc acctcctccg gccggaacag | 600 |
| 10 | gaaagcagct ccaggcagct ggggtccacaa aaatctcgt tccctgaggt ctcagaggca | 660 |
| | gtggcccagg agcatctggt caccttcggg aaaaaccggc ttggcaaagg ctcccccgag | 720 |
| | ggcacgcgtt tcccggacag tgaggcagga cctaaactct tccgttaaca ctacattttt | 780 |
| | cgcatttctg cagtgtttgc actctcaggc cccaccattt ccccgcatct cttagggaga | 840 |
| | agttctcgac gtcccacctc ccctggaagg gtgctgctcc cagagacctt caggccaatg | 900 |
| 15 | gcccattctc agtgccctca ggggagaggg ggggtgcagaa aaacagcctg ggtcacaaaa | 960 |
| | gaggtgcgag ggctgtgaga tcccggaggc accgacggga agcgagacgg agaacaggag | 1020 |
| | ggcaggacgg gctggagggt ggggatactg cagatggagg gagccacggt gggggagggc | 1080 |
| | gtggacctga ccgtcctggc acaaggcgggt cgggtgcaga cctccaggcc ctccgggtta | 1140 |
| | aggtgccgcc cagagccctc aggccggggg cgcacggaaa ccacaggcag ggtgcgcgtg | 1200 |
| 20 | gagggacggg gaaagcgggg cggtgtgggg aaggcgcccc gggaacctga acctcccacc | 1260 |
| | ccgcctcagt ctcgaccact ccttaagccc caccocggcc caggtaaggc gcagtccacc | 1320 |
| | cccattccca gtagattaac gcacagggtg gggcgcgctc gggacatagc tgcgctaggg | 1380 |
| | gacagcgcg cccagcccagt cgcggggggc agggagcagg cggggcccag caggaaccca | 1440 |
| | gctttgttag cgatgctccc cgtgagccac gcgccacgg tacgcgcttc ctcaatgggg | 1500 |
| 25 | ccgggcgtgg agccgcgccc tgcgcgattg gccaacggg tggccacga ttggctgaga | 1560 |
| | ccctggcccc cgctcctcg gcccaggag ggtggggcgt ggggtgtggc tgcgcggcgc | 1620 |
| | gtgctgcccc cggggatctt gcgcgcctcc cgaacagccg tgttgtcgcc agggccgcgc | 1680 |
| | cttccctccc acagcgcgcg ctgcgcgtgc gaaggctctg cggtccttg gactggcggg | 1740 |
| | gctgcgcgcg gggttagggt ggggttacgg gaaggctcaa cccaggacct gcgtaccttg | 1800 |
| 30 | ctttgggggc gcactaagca cctgccggga gcagggggcg caccgggaac tcgcagattt | 1860 |
| | cgccagttgg gcgcactgg gatctgtgga ctgcgtccgg gggatgggct agggggacat | 1920 |
| | gcgcacgctt tgggccttac agaattgat cgcgcgagg ggagggcgaa gcgtggcggg | 1980 |
| | agggcgaggc gaaggaaagga gggcgtgaga aaggcgacgg cggcggcgcg gaggagggtt | 2040 |
| | atctatacat ttaaaaacca gccgcctcgc ccgcgcctgc ggagacctg gagagtccg | 2100 |
| 35 | ccgcacgcgc gggacacgag cgtcccacgc tccctggcgc gtacggcctg ccaccactag | 2160 |
| | gcctcctatc cccgggctcc agacgaccta ggacgcgtgc cctggggagt tgctggcgg | 2220 |

cgccgtgccga gaagccccct tggggcgcca cagttttccc cgtcgccctcc ggttcctctg 2280
cctgcacctt cctgcggcgc gccgggacct ggagcgggcg ggtggatgca ggcgcgatgg 2340
acggcggcac actgcccagg tccgcgcccc ctgcgcccc cgtccctgtc ggctgcgtg 2400
cccggcggag acccgcgctc ccggaactgt tgcgtgcag ccggcggcgg cgaccggcca 2460
5 ccgcagagac cggaggcggc gcagcggccg tagcgggcg c 2501

<210> 81

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 81

aatcctccaa attctaaaaa ca

22

<210> 82

20 <211> 20

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 82

aggaaaggga gtgagaaaat

20

30

<210> 83

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 83

5 ggataggagt tgggattaag at

22

<210> 84

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 84

aaatcttttt caacaccaaa at

22

<210> 85

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 85

aaccctttct tcaaattaca aa

22

30

<210> 86

<211> 21

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 86

5 tgattgggtt ttagggaaat a

21

<210> 87

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 87

ttgaaaataa gaaaggttga gg

22

<210> 88

20 <211> 19

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 88

cttctacccc aaatcccta

19

30

<210> 89

<211> 18

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 89

5 tgtttgggat tgggtagg

18

<210> 90

<211> 23

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 90

cataaccttt acctatctcc tca

23

<210> 91

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 91

tttttagattg aggttttagg gt

22

30

<210> 92

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 92

5 atccattcta cctccttttt ct

22

<210> 93

<211> 18

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 93

ggaggggaga gggttatg

18

<210> 94

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 94

tactatacac accccaaaac aa

22

30

<210> 95

<211> 19

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 95

5 ttttgggaat gggttgtat

19

<210> 96

<211> 21

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 96

ctacccttaa cctccatcct a

21

<210> 97

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 97

ttgttgggag tttttaagtt tt

22

30

<210> 98

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 98

5 caaattctcc ttccaaataa at

22

<210> 99

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 99

gtaatttgaa gaaagttgag gg

22

<210> 100

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 100

ccaacaacta aacaaaacct ct

22

30

<210> 101

<211> 20

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 101

5 ggagttgtat tgttgggaga

20

<210> 102

<211> 21

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 102

taaaacccca attttcacta a

21

<210> 103

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 103

tttgtattag gttggaagtg gt

22

30

<210> 104

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 104

5 cccaaataaaa tcaacaacaa ca

22

<210> 105

<211> 22

<212> DNA

10 <213> Artificial Sequence

 $\langle 220 \rangle$

<223> primer

15 <400> 105

gatttttggg gaggaagtta ag

22

<210> 106

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 106

aaaactaaaa accaaacca ta

22

30

<210> 107

<211> 20

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 107

5 tgggggttagt ttaggatagg

20

<210> 108

<211> 25

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 108

cttaaaaaca ctaaaacttc tcaaa

25

<210> 109

20 <211> 21

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 109

tttttgtatt ggggtaggtt t

21

30

<210> 110

<211> 24

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 110

5 cccaactatc tctctcctct ataa 24

<210> 111

<211> 25

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 111

attagaagtg aaagtaatgg aattt 25

<210> 112

20 <211> 19

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 112

tcaatttcca aaaaccaac 19

30 <210> 113

<211> 22

<212> DNA

<213> Artificial Sequence

35 <220>

<223> primer

<400> 113

5 gggatggggtt attagttgta aa

22

<210> 114

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 114

ccttcacaca aaactacaaa aa

22

<210> 115

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 115

taattgaagg ggttaatagt gg

22

30

<210> 116

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 116

5 aaaaccaaaa ccaaaactaa aa

22

<210> 117

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 117

agtggaattg gagtttagat gt

22

<210> 118

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 118

aacaaaataa aaacttctcc ca

22

30

<210> 119

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 119

5 taggggaaaa gttagagttg ag

22

<210> 120

<211> 18

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 120

cccattaacc cacaaaaa

18

<210> 121

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 121

at tttagttt gtgaaatggg at

22

30

<210> 122

<211> 21

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 122

5 tcttaaccaa taaccctca c 21

<210> 123

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 123

gtggggttttg ggtagttata ga 22

<210> 124

20 <211> 20

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 124

taacctcctc tccttaccaa 20

30 <210> 125

<211> 22

<212> DNA

<213> Artificial Sequence

35 <220>

<223> primer

<400> 125

5 taggatgggg agagtaatgt tt 22

<210> 126

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 126

acaacttatc caacttccat tc 22

<210> 127

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 127

tcccacaaaa actaaacaat ta 22

30 <210> 128

<211> 21

<212> DNA

<213> Artificial Sequence

35 <220>

<223> primer

<400> 128

5 aggttttaga tgaaggggtt t

21

<210> 129

<211> 23

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 129

tttggagggt ttagtagaag tta

23

<210> 130

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 130

cccaataatc acaaaataaa ca

22

30

<210> 131

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 131

5 atacaacctc aaatcctatc ca

22

<210> 132

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 132

agggagaagg aagttatttg tt

22

<210> 133

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 133

ggaagatgag gaagttgatt ag

22

30

<210> 134

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 134

5 cctacaaccc tatcctctaa aa 22

<210> 135

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 135

ttagtagggg tgtgagtgtt tt 22

<210> 136

20 <211> 23

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 136

caaacaaaac ttctatctca acc 23

30 <210> 137

<211> 21

<212> DNA

<213> Artificial Sequence

35 <220>

<223> primer

<400> 137

5 ttataggggt gagtttggga t

21

<210> 138

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 138

taaacaaaca acaaattctt ca

22

<210> 139

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 139

tgaaaatgaa ggtatggagt tt

22

30

<210> 140

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 140

5 ttaaaacccat ataatccctc ca

22

<210> 141

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 141

tatgttttgt tttgttttga ga

22

<210> 142

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 142

aaccccatca cttttatttc tt

22

30

<210> 143

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 143

5 ggggtgtagaa gtgttttaggt tt

22

<210> 144

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 144

tttctccct tacaacaata ac

22

<210> 145

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 145

tccccttcca actatatctc tc

22

30

<210> 146

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 146

5 tgagagtgtt ttagggaagt tt 22

<210> 147

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 147

aaaacccaaa cataaaccaa aa 22

<210> 148

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 148

gattaggagg gtttgttgag at 22

30

<210> 149

<211> 21

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 149

5 aatggttgat gattttggtt t

21

<210> 150

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 150

actctcttcc ctataccct aa

22

<210> 151

20 <211> 24

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 151

tgtagtaga gtttagga ggtt

24

30

<210> 152

<211> 22

<212> DNA

<213> Artificial Sequence

35

<220>

<223> primer

<400> 152

5 acactaccta tccttaccac ac 22

<210> 153

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 153

tttttgtttt tatggggtgt at 22

<210> 154

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 154

ttaaatatcc cttccttaac ca 22

30 <210> 155

<211> 23

<212> DNA

<213> Artificial Sequence

35 <220>

<223> primer

<400> 155

5 agttagaaga ggagttagga tgg

23

<210> 156

<211> 22

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 156

taattttcca ataccattt tc

22

<210> 157

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 157

30 tgggtagtat tttgttggt tt

22

<210> 158

<211> 22

<212> DNA

35 <213> Artificial Sequence

<220>

<223> primer

<400> 158

5 cctaaaaact ctctcatcct ca

22

<210> 159

<211> 23

<212> DNA

10 <213> Artificial Sequence

<220>

<223> primer

15 <400> 159

agtggtttag gagtatttgg tta

23

<210> 160

20 <211> 22

<212> DNA

<213> Artificial Sequence

<220>

25 <223> primer

<400> 160

aactccctcc atctacaata tc

22

30

Figure 1

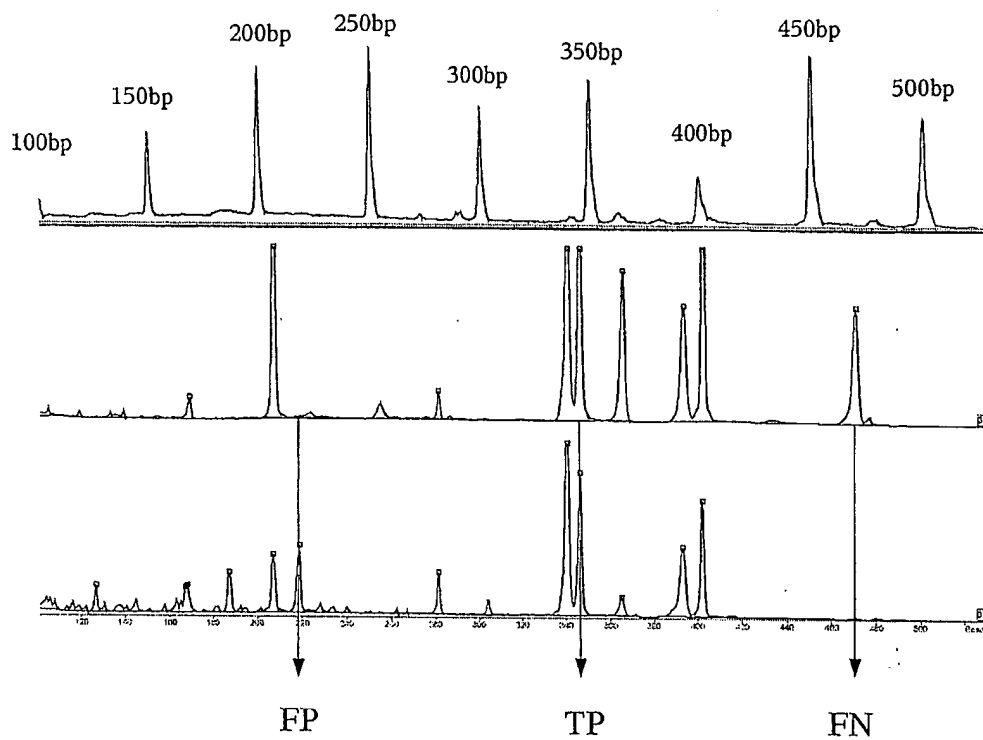


Figure 2

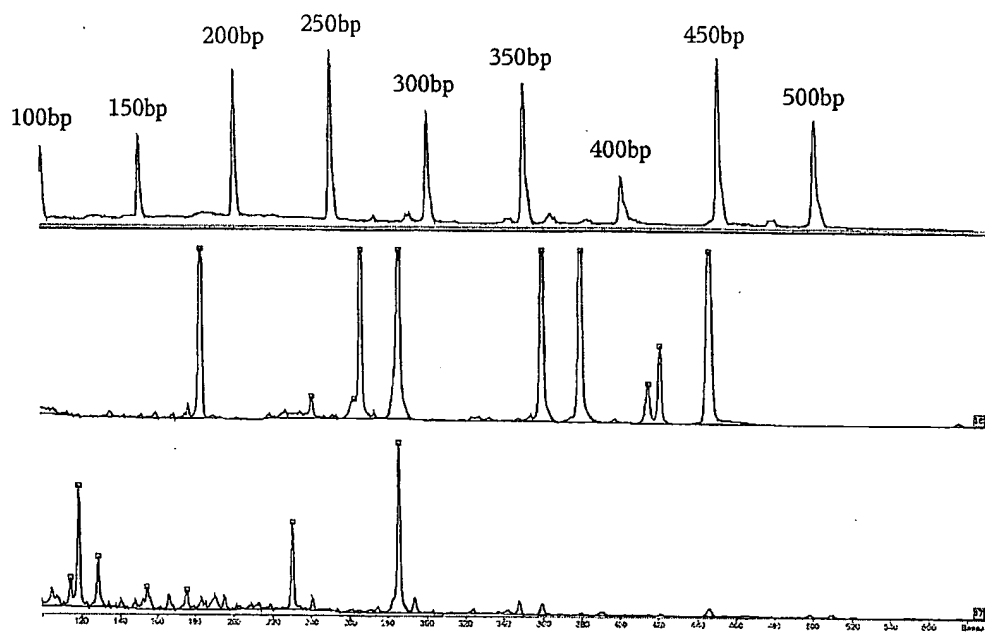


Figure 3

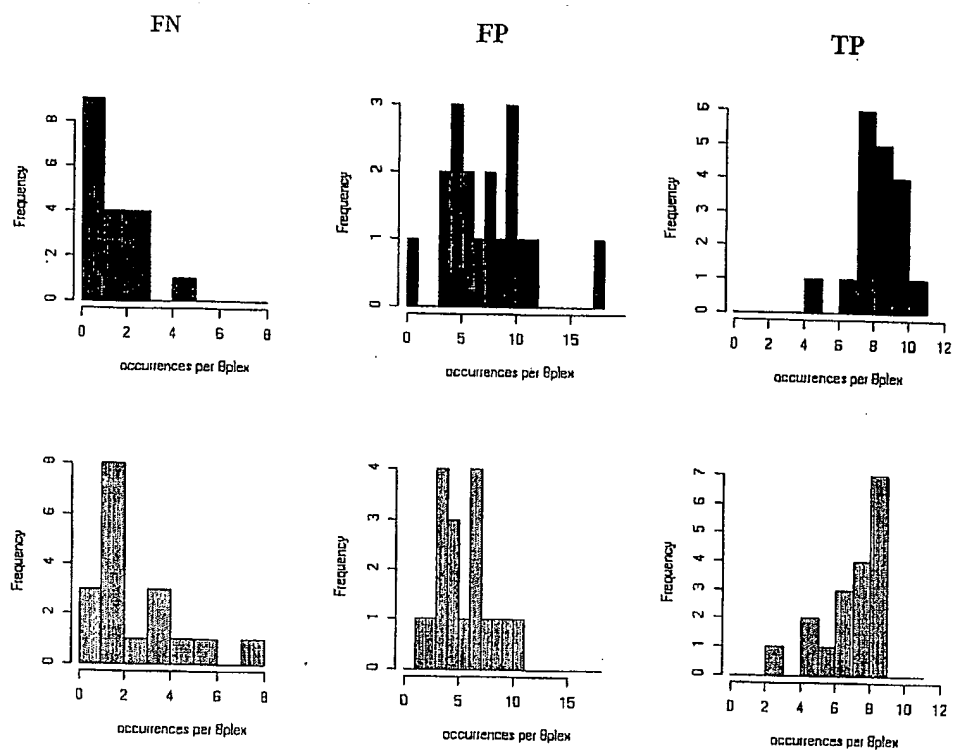


Figure 4

| A | | | |
|------|----|----|----|
| | TP | FP | FN |
| 11-a | 9 | 8 | 0 |
| 11-b | 9 | 7 | 0 |
| 12 | 7 | 9 | 0 |
| 13 | 7 | 6 | 1 |
| 14 | 9 | 11 | 1 |
| 15 | 7 | 5 | 1 |
| 21-a | 8 | 4 | 2 |
| 21-b | 8 | 4 | 0 |
| 22 | 8 | 9 | 2 |
| 23-a | 7 | 5 | 0 |
| 23-b | 7 | 10 | 1 |
| 24 | 10 | 17 | 0 |
| 25 | 8 | 0 | 0 |
| 31 | 9 | 3 | 0 |
| 32 | 4 | 7 | 4 |
| 33 | 6 | 9 | 2 |
| 34 | 8 | 3 | 0 |
| 35 | 7 | 4 | 2 |

| | TP | FP | FN |
|-------|------|------|------|
| Ø | 7,67 | 6,72 | 0,89 |
| STABW | 1,37 | 3,86 | 1,13 |

| C | | | |
|------|----|----|----|
| | TP | FP | FN |
| 11-a | 8 | 7 | 1 |
| 11-b | 8 | 4 | 1 |
| 12 | 7 | 6 | 1 |
| 13 | 8 | 6 | 0 |
| 14 | 6 | 6 | 2 |
| 15 | 4 | 9 | 5 |
| 21-a | 7 | 3 | 0 |
| 21-b | 8 | 4 | 0 |
| 22 | 4 | 6 | 3 |
| 23-a | 6 | 4 | 3 |
| 23-b | 8 | 5 | 1 |
| 24 | 8 | 10 | 4 |
| 25 | 5 | 1 | 3 |
| 31 | 7 | 2 | 1 |
| 32 | 8 | 3 | 1 |
| 33 | 2 | 3 | 7 |
| 34 | 7 | 8 | 1 |
| 35 | 6 | 3 | 1 |

| | TP | FP | FN |
|---|------|------|------|
| Ø | 6,5 | 5 | 1,94 |
| | 1,76 | 2,45 | 1,89 |

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/08602

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, EMBASE, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--|-----------------------|
| X | WO 01 62960 A (BERLIN KURT ;EPIGENOMICS AG (DE); OLEK ALEXANDER (DE)) 30 August 2001 (2001-08-30) page 10, line 9 - line 12; claims 1-26 page 13, line 10 - line 29 --- | 1-51 |
| X | HERMAN J G ET AL: "METHYLATION-SPECIFIC PCR: A NOVEL PCR ASSAY FOR METHYLATION STATUS OF CPG ISLANDS" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 93, 1 September 1996 (1996-09-01), pages 9821-9826, XP002910406 ISSN: 0027-8424 cited in the application page 9821 -page 9823 --- -/-- | 1-51 |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *G* document member of the same patent family

Date of the actual completion of the international search

10 December 2003

Date of mailing of the international search report

02/01/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Favre, N

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 03/08602

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--|-----------------------|
| X | US 5 786 146 A (HERMAN JAMES G ET AL) 28 July 1998 (1998-07-28) page 6, line 1 -page 7, line 6; claim 1 --- | 1-51 |
| X | WO 01 44504 A (FOX JAYNE CATHERINE ;HAQUE KEMAL (GB); LITTLE STEPHEN (GB); ASTRAZ) 21 June 2001 (2001-06-21) page 6, line 26 -page 8, line 7; figure 1 --- | 1-51 |
| X | WO 98 56952 A (UNIV SOUTHERN CALIFORNIA) 17 December 1998 (1998-12-17) page 9, line 20 -page 10, line 2; claims 9,10 --- | 1-51 |
| X | REIN ET AL: "Identifying 5-methylcytosine and related modifications in DNA genomes" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 26, no. 10, 1998, pages 2255-2264, XP002143106 ISSN: 0305-1048 cited in the application abstract; figure 2 --- | 1-51 |
| X | WO 00 70090 A (UNIV SOUTHERN CALIFORNIA) 23 November 2000 (2000-11-23) page 13, line 29 -page 14, line 34; claim 1 --- | 1-51 |
| X | ZESCHNIGK M ET AL: "A SINGLE-TUBE PCR TEST FOR THE DIAGNOSIS OF ANGELMAN AND PRADER-WILLI SYNDROME BASED ON ALLELIC METHYLATION DIFFERENCES AT THE SNRPN LOCUS" EUROPEAN JOURNAL OF HUMAN GENETICS, KARGER, BASEL, CH, vol. 5, no. 2, 1997, pages 94-98, XP009011533 ISSN: 1018-4813 cited in the application abstract; figure 1 --- | 1-51 |
| A | GRIFFIN HG AND GRIFFIN AM (EDS.): "PCR TECHNOLOGY: Current Innovations" 1994 , CRC PRESS , BOCA RATON XP008025537 page 5 -page 11 --- | 1-51 |
| A | US 6 007 231 A (BISHOP ROBERT ET AL) 28 December 1999 (1999-12-28) the whole document --- | 1-51 |
| | -/-- | |

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/08602

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|---|-----------------------|
| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| A | SEIJEN A M ET AL: "Systematic design of mouse Vh gene family-specific oligonucleotides" JOURNAL OF IMMUNOLOGICAL METHODS, ELSEVIER SCIENCE PUBLISHERS B.V., AMSTERDAM, NL, vol. 254, no. 1-2, 1 August 2001 (2001-08-01), pages 161-168, XP004245450 ISSN: 0022-1759 the whole document --- | 1-51 |
| A | WO 00 49177 A (DU PONT ; ROUVIERE PIERRE (US)) 24 August 2000 (2000-08-24) the whole document --- | 1-51 |
| A | SCHULER G D: "SEQUENCE MAPPING BY ELECTRONIC PCR" GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, US, vol. 7, 1997, pages 541-550, XP000872176 ISSN: 1088-9051 cited in the application the whole document --- | 1-51 |
| A | ENGELS W R: "Contributing software to the internet: the Amplify program" TRENDS BIOCHEM. SCI., vol. 18, 1993, pages 448-450, XP001161002 the whole document --- | 1-51 |
| P, X | US 2003/068625 A1 (BARRETT WADE A ET AL) 10 April 2003 (2003-04-10) claims 1-32; figures 1-4 ----- | 39-51 |

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 03/08602

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Insofar as they include the step of isolating a nucleic sample, i.e. a step which could be performed on the living human or animal body, claims 1-38 relate to a method of treatment and/or diagnostic performed on the living human or animal body, the search has been carried out and restricted on in vitro methods.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 03/08602

| Patent document cited in search report | | Publication date | Patent family member(s) | Publication date |
|---|----|---------------------|----------------------------|---------------------|
| WO 0162960 | A | 30-08-2001 | DE 10010282 A1 | 06-09-2001 |
| | | | AU 4227701 A | 03-09-2001 |
| | | | CA 2401233 A1 | 30-08-2001 |
| | | | WO 0162960 A2 | 30-08-2001 |
| | | | EP 1257670 A2 | 20-11-2002 |
| | | | JP 2003525041 T | 26-08-2003 |
| | | | US 2003157510 A1 | 21-08-2003 |
| US 5786146 | A | 28-07-1998 | CA 2257104 A1 | 11-12-1997 |
| | | | EP 0954608 A1 | 10-11-1999 |
| | | | IL 127342 A | 25-07-2002 |
| | | | JP 2000511776 T | 12-09-2000 |
| | | | WO 9746705 A1 | 11-12-1997 |
| | | | US 6200756 B1 | 13-03-2001 |
| | | | US 6265171 B1 | 24-07-2001 |
| | | | US 6017704 A | 25-01-2000 |
| WO 0144504 | A | 21-06-2001 | AU 2194801 A | 25-06-2001 |
| | | | WO 0144504 A2 | 21-06-2001 |
| WO 9856952 | A | 17-12-1998 | AU 7829398 A | 30-12-1998 |
| | | | US 6251594 B1 | 26-06-2001 |
| | | | WO 9856952 A1 | 17-12-1998 |
| | | | US 2002177154 A1 | 28-11-2002 |
| | | | US 2003211473 A1 | 13-11-2003 |
| WO 0070090 | A | 23-11-2000 | US 6331393 B1 | 18-12-2001 |
| | | | AU 4712200 A | 05-12-2000 |
| | | | CA 2372665 A1 | 23-11-2000 |
| | | | EP 1185695 A1 | 13-03-2002 |
| | | | JP 2002543852 T | 24-12-2002 |
| | | | WO 0070090 A1 | 23-11-2000 |
| | | | US 2002086324 A1 | 04-07-2002 |
| US 6007231 | A | 28-12-1999 | AU 3632097 A | 06-03-1998 |
| | | | CA 2263731 A1 | 19-02-1998 |
| | | | EP 0925372 A1 | 30-06-1999 |
| | | | WO 9806872 A1 | 19-02-1998 |
| WO 0049177 | A | 24-08-2000 | CA 2359645 A1 | 24-08-2000 |
| | | | EP 1153141 A2 | 14-11-2001 |
| | | | WO 0049177 A2 | 24-08-2000 |
| | | | US 2003113886 A1 | 19-06-2003 |
| | | | US 6365376 B1 | 02-04-2002 |
| | | | US 2002127666 A1 | 12-09-2002 |
| US 2003068625 | A1 | 10-04-2003 | WO 03021259 A1 | 13-03-2003 |
| | | | US 2003108919 A1 | 12-06-2003 |
| | | | US 2003073093 A1 | 17-04-2003 |